

# Raman resonance in iron-based superconductors: The magnetic scenario

Alberto Hinojosa,<sup>1</sup> Jiashen Cai,<sup>1</sup> and Andrey V. Chubukov<sup>1</sup>

<sup>1</sup>*Department of Physics, University of Minnesota, Minneapolis, Minnesota 55455, USA*

We perform theoretical analysis of polarization-sensitive Raman spectroscopy on  $\text{NaFe}_{1-x}\text{Co}_x\text{As}$ ,  $\text{EuFe}_2\text{As}_2$ ,  $\text{SrFe}_2\text{As}_2$ , and  $\text{Ba}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$ , focusing on two features seen in the  $B_{1g}$  symmetry channel (in one Fe unit cell notation): the strong temperature dependence of the static, uniform Raman response in the normal state and the existence of a collective mode in the superconducting state. We show that both features can be explained by the coupling of fermions to pairs of magnetic fluctuations via the Aslamazov-Larkin process. We first analyze magnetically-mediated Raman intensity at the leading two-loop order and then include interactions between pairs of magnetic fluctuations. We show that the full Raman intensity in the  $B_{1g}$  channel can be viewed as the result of the coupling of light to Ising-nematic susceptibility via Aslamazov-Larkin process. We argue that the singular temperature dependence in the normal state is the combination of the temperature dependencies of the Aslamazov-Larkin vertex and of Ising-nematic susceptibility. We discuss two scenarios for the resonance below  $T_c$ . In one, the resonance is due to the development of a pole in the fully renormalized Ising-nematic susceptibility. The other is the orbital excitonic scenario, in which spin fluctuations generate an attractive interaction between low-energy fermions.

## I. INTRODUCTION

Raman scattering in Fe-based superconductors has attracted substantial interest in the past few years due to the number of new features associated with multi-orbital/multi-band nature of these materials (see, e.g., Refs. [1–8]). The subject of this paper is the theoretical analysis of the features in Raman scattering revealed by polarization-sensitive Raman spectroscopy in the normal and the superconducting states of Fe-based materials  $\text{NaFe}_{1-x}\text{Co}_x\text{As}$  [9],  $\text{AFe}_2\text{As}_2$ ,  $\text{A} = \text{Eu}, \text{Sr}$  [10 and 11], and  $\text{Ba}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$  [12 and 13]. Polarized light was used to probe the Raman response in different symmetry channels, classified by the irreducible representations of the  $D_{4h}$  crystallographic point group [1]. In the normal state the real part of the (almost) static and uniform Raman susceptibility in the  $B_{1g}$  channel in one-iron unit cell notation (same as the  $B_{2g}$  channel in two-iron notation used in Refs. [9, 10, and 12]) is strongly temperature dependent—it increases below 300 K roughly as  $1/(T - T_0)$ , where  $T_0$  is positive at small doping, but changes sign and becomes negative above optimal doping. In the superconducting state the imaginary part of the  $B_{1g}$  Raman susceptibility displays a strong resonance-type peak at around  $50 \text{ cm}^{-1}$ . There is no such resonance peak in other channels, although the Raman intensity in the  $A_{1g}$  channel does show a broad maximum at a somewhat higher frequency [9].

In the effective mass approximation (in which the coupling of light to fermions is proportional to the square of vector potential) the measured Raman intensity in a particular channel ( $A_{1g}$ ,  $B_{1g}$ ,  $B_{2g}$ , ...) is proportional to the imaginary part of the fully renormalized particle-hole polarization bubble  $\chi_R(\mathbf{p}, \Omega)$  with proper symmetry factors in the vertices, taken at vanishingly small transferred momentum  $\mathbf{p}$  and finite transferred frequency  $\Omega$  [15–17].

The free-fermion polarization bubble vanishes in the normal state for  $\Omega > v_F p$ , where  $v_F$  is the Fermi velocity, and obviously it cannot account for the observed

strong temperature dependence of  $B_{1g}$  Raman intensity above  $T_c$ . It is nonzero in the superconducting state, but does not display a peak. The effect must then come from the renormalization of the Raman vertex due to coupling to some low-energy fluctuations (final state interaction in Raman literature [18]). This coupling may come from three different sources. First, the  $B_{1g}$  Raman vertex changes sign under  $k_x \leftrightarrow k_y$  (i.e., under interchanging the  $x$  and  $y$  directions in real space), hence it couples to strain (structural fluctuations). Second, the  $B_{1g}$  vertex is anti-symmetric with respect to the interchange between the iron  $d_{xz}$  and  $d_{yz}$  orbitals and hence couples to orbital fluctuations. Third, symmetry allows the coupling between the  $B_{1g}$  Raman vertex and Ising-nematic spin fluctuations [the ones that distinguish between the magnitudes of spin-density-wave order parameters with ordering vectors  $(0, \pi)$  and  $(\pi, 0)$ ] because both are anti-symmetric with respect to  $90^\circ$  rotations in the momentum space.

Structural fluctuations, orbital fluctuations, and Ising-nematic spin-fluctuations are the three key candidates to drive the nematic order, observed in most of the Fe-based materials. How to choose the primary candidate among these three has become one of the key issues in the studies of Fe-based superconductors [19]. We intend to verify whether the Raman data can help distinguish between the three candidates.

The effects of structural and orbital fluctuations has been discussed before (see Refs. [20], [21] and references therein). Structural fluctuations (acoustic phonons associated with strain) practically do not affect the Raman intensity because the coupling to phonons changes the  $B_{1g}$  Raman susceptibility  $\chi_R(\mathbf{p}, \Omega)$  to

$$\tilde{\chi}_R(\mathbf{p}, \Omega) = \left[ \chi_R^{-1}(\mathbf{p}, \Omega) - \frac{\lambda_{ph}^2 p^2}{C_{ph}^2 p^2 - \Omega^2} \right]^{-1}, \quad (1)$$

where  $\lambda_{ph}$  is electron-phonon coupling and  $C_{ph}$  is the elastic constant for orthorhombic strain. Such coupling is

relevant in the static limit, where  $\tilde{\chi}_R^{-1}(\mathbf{p}, 0) = \chi_R^{-1}(\mathbf{p}, 0) - (\lambda_{ph}/C_{ph})^2$  differs from  $\chi_R^{-1}(\mathbf{p}, 0)$  by a constant term, but is irrelevant in the limit of vanishing  $p$  and finite  $\Omega$ , where Raman measurements have been performed [9, 12, and 13]. (In the  $B_{1g}$  channel the minimum  $p$  is, strictly speaking, nonzero [22], but is generally of order of inverse system size).

Orbital fluctuations do affect the  $B_{1g}$  Raman susceptibility via renormalizations involving particular combinations of intra-orbital and inter-orbital Hubbard and Hund terms, compatible with the fact that the  $B_{1g}$  Raman vertex changes sign between  $d_{xz}$  and  $d_{yz}$  orbitals. By orbital fluctuations we mean fluctuations which renormalize  $B_{1g}$  Raman vertex by inserting series of ladder and bubble diagrams, as shown schematically in Fig. 1 and in more detail in Figs. 17 and 18 in the Appendix.

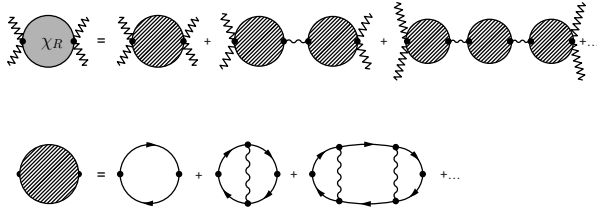


Figure 1. Ladder and bubble diagrams for the renormalization of the  $B_{1g}$  Raman intensity within RPA. Each combination of Green's functions with equal momenta gives a particle-hole polarization bubble  $\Pi_{B_{1g}}(\Omega)$  and the combination of vertical and horizontal interaction lines gives the coupling  $\lambda$ . We obtain  $\lambda$  explicitly in the Appendix for a system with dominant intra-orbital Hubbard interaction.

In the band basis, the interaction lines in these diagrams are Hubbard and Hund terms dressed by coherence factors, associated with the transformation from the orbital to the band basis, and projected into  $B_{1g}$  channel. (Each interaction contains one incoming and one outgoing fermion with momentum  $\mathbf{k}$  and one incoming and one outgoing fermion with momentum  $\mathbf{p}$ , and its  $B_{1g}$  component is proportional to  $\cos 2\theta_k \cos 2\theta_p$ , where  $\tan \theta_k = k_y/k_x$ ). In the approximation when only ladder and bubble diagrams are kept [often called random-phase approximation (RPA)], the Raman response in the  $B_{1g}$  channel is given by (see, e.g., Ref. [23]),

$$\chi_R(\Omega) = \frac{\Pi_{B_{1g}}(\Omega)}{1 + \lambda \Pi_{B_{1g}}(\Omega)}, \quad (2)$$

where  $\chi_R(\Omega) = \chi_R(p \rightarrow 0, \Omega)$ ,  $\lambda$  is the proper combination of interactions in  $B_{1g}$  channel, and  $\Pi_{B_{1g}}(\Omega)$  is the particle-hole polarization function at zero momentum transfer, summed over all bands with the  $B_{1g}$  form factor. [We define  $\Pi_{B_{1g}}(\Omega)$  as  $i \int d^2k d\nu / (2\pi)^3 \times G(\mathbf{k}, \nu) G(\mathbf{k}, \nu + \Omega)$ . With this sign convention,  $\text{Re } \Pi_{B_{1g}}(\Omega)$  in a superconductor is positive at  $\Omega < 2\Delta$ , where  $\Delta$  is the superconducting gap]. In the language of Fermi-liquid theory,  $\lambda \Pi_{B_{1g}}(\Omega = 0)$  is the  $B_{1g}$  component of the quasiparticle interaction function, and  $\lambda \Pi_{B_{1g}}(\Omega = 0) = -1$  would correspond to  $B_{1g}$  Pomeranchuk instability.

Because free-fermion  $\Pi_{B_{1g}}(\Omega)$  vanishes in the normal state (the poles of both Green's functions are in the same frequency half plane), Eq. (2) cannot explain the normal state behavior of the Raman response. However,  $\chi_R(\Omega)$  from Eq. (2) is nonzero in a superconductor, because  $\Pi_{B_{1g}}(\Omega)$  becomes nonzero, and for negative  $\lambda$  it displays a resonance peak. The resonance develops by the same reason as the excitonic spin resonance in a  $d$ -wave superconductor [24]: the imaginary part of the polarization function  $\Pi(\Omega)$  vanishes for  $\Omega < 2\Delta$ , while the real part of  $\Pi(\Omega)$  is positive and diverges at  $\Omega = 2\Delta$ . As the result, for negative  $\lambda$ , the denominator in (2) is guaranteed to pass through zero at some frequency below  $2\Delta$ , and a sharp resonance in  $\chi_R(\Omega)$  appears at this frequency [23].

This would be the most natural explanation of the Raman resonance. The problem, however, is how to justify that  $\lambda$  is negative, i.e., that there is an attraction in the  $B_{1g}$  ( $d$ -wave) charge Pomeranchuk channel. If intra-orbital Hubbard repulsion is the dominant interaction term,  $\lambda$  is definitely positive and orbital fluctuations do not give rise to the resonance in the Raman intensity (we show this in the Appendix). The coupling  $\lambda$  does become negative when inter-orbital interaction  $U'$  and exchange Hund interaction  $J$  are included and  $U'$  is set to be about equal to  $U$  and larger than  $J$  [25]. However, the relation  $U' \approx U$  gets broken once one starts integrating out high-energy fermionic excitations [26] or includes lattice effects [27]. In a generic case it is natural to expect that the intra-orbital Hubbard interaction is the strongest interaction between Fe  $d$ -orbitals. If so, the coupling  $\lambda$  is positive and there is no resonance in  $\chi_R(\Omega)$  within RPA.

In this paper we analyze whether the increase of  $\chi_R(\Omega)$  in the  $B_{1g}$  channel in the normal state and the sharp peak in the Raman response in this channel below  $T_c$  can be due to Ising-nematic spin fluctuations associated with stripe magnetism. The advantage of the magnetic scenario is that Ising-nematic fluctuations are enhanced even when intra-orbital Hubbard interaction is the dominant interaction between fermions. The only requirement is that the magnetic order should be stripe rather than checkerboard [28].

The coupling of the Raman vertex to a pair of spin fluctuations with momenta near  $\mathbf{Q} = (0, \pi)$  [or  $(\pi, 0)$ ] occurs via the Aslamazov-Larkin (AL) process. The corresponding diagram is presented in Fig. 2. AL diagrams for Raman scattering have been earlier discussed in Ref. [29], but in a different context. The lowest-order AL type diagram (the one shown in Fig. 2) contains two triangular vertices made out of fermionic Green's functions from hole and electron bands, and two spin-fluctuation propagators. The vertex between fermions and spin fluctuations can be obtained by decomposing the antisymmetrized interaction into spin and charge parts, by focusing on spin-spin part, and by using the Hubbard-Stratonovich method to transform the interaction between fermionic spins into spin-spin interaction between a fermion and a collective bosonic variable in the spin channel [28].

We show that in the normal state, above a certain temperature, each triangular vertex  $\Gamma_{tr}$  scales as  $1/T$ , while the convolution of the two spin propagators at equal frequencies (i.e.,  $T \sum_{\nu_n} \int d^2\mathbf{q}/(2\pi)^2 \times [\chi^s(\mathbf{Q} + \mathbf{q}, \nu_n)]^2$ ) scales as  $T$ . As the consequence, the Raman susceptibility from Fig. 2 scales as  $1/T$ . This holds in both  $A_{1g}$  and  $B_{1g}$  channels. Higher-order processes, shown in Fig. 3 and in more detail in Fig. 6 below, however, distinguish between  $A_{1g}$  and  $B_{1g}$  Raman responses. Namely, an attractive interaction between magnetic fluctuations in the  $B_{1g}$  channel increases  $B_{1g}$  Raman response and changes  $1/T$  dependence into  $1/(T - T_0)$  (see Refs. [30, 31] and Sec.III below), while the (much stronger) repulsive interaction in the  $A_{1g}$  channel almost completely eliminates the temperature dependence of  $A_{1g}$  Raman response. This behavior fully agrees with the data.

In the superconducting state, the  $1/T$  behavior of  $\Gamma_{tr}$  is cut by the gap opening, while  $\chi^2(\Omega) = \int d\nu d^2\mathbf{q}/(2\pi)^3 \times \chi^s(\mathbf{q} + \mathbf{Q}, \nu) \chi^s(\mathbf{q} + \mathbf{Q}, \nu + \Omega)$  becomes singular. The real part of  $\chi^2(\Omega)$  diverges at  $\Omega = 2\Omega_{mag}$ , where  $\Omega_{mag}$  is the minimal frequency of the magnetic resonance in the superconducting state, and its imaginary part jumps at this frequency from zero to a finite value. Higher-order terms change  $\chi^2(\Omega)$  into  $\chi_{I-nem}(\Omega) = \chi^2(\Omega)/[1 + 2g\chi^2(\Omega)]$  (see Sec. III), where  $g$  is negative (attractive) when magnetic order is of stripe type [28]. Approximating the triangular vertex  $\Gamma_{tr}$  by a constant we then obtain

$$\chi_R(\Omega) = \Gamma_{tr}^2 \chi_{I-nem}(\Omega) = \Gamma_{tr}^2 \frac{\chi^2(\Omega)}{1 + 2g\chi^2(\Omega)}. \quad (3)$$

The combination of  $g < 0$  and the fact that the real part of  $\chi^2(\Omega)$  is positive and diverges at  $\Omega = 2\Omega_{mag}$  guarantees that  $1 + 2g\chi^2(\Omega)$  passes through zero at some  $\Omega = \Omega_{res,1} < 2\Omega_{mag}$ . At this frequency the Raman intensity  $\text{Im } \chi_R(\Omega)$  displays a  $\delta$ -functional peak.

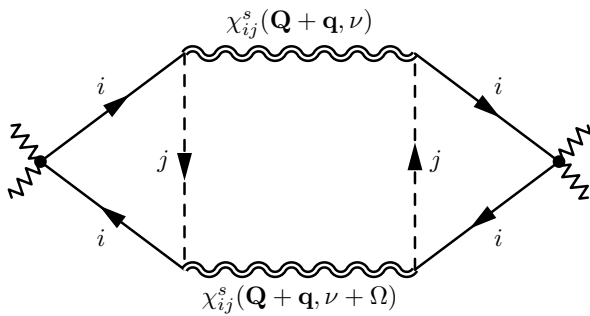


Figure 2. The lowest order (two-loop) AL diagram for the Raman intensity. The momenta  $\mathbf{Q}_1 = (\pi, 0)$  and  $\mathbf{Q}_2 = (0, \pi)$ . The solid and dashed lines represent fermions from different bands with band indices  $i$  and  $j$ . The sinuous lines represent spin fluctuations and the external jagged lines are the coupling to photons.

We show that approximating  $\Gamma_{tr}$  by a constant is justified if typical fermionic frequencies in the triangular diagram for  $\Gamma_{tr}$  are larger than  $\Omega_{mag}$ . These relevant

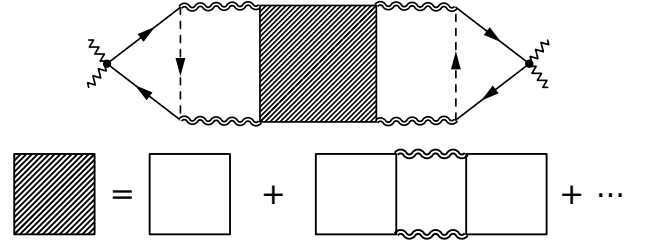


Figure 3. Schematic representation of higher-order contributions to the Raman response. We show higher-order terms in more detail in Fig. 6.

frequencies are of order  $\Delta$ , hence the analysis is justified when  $\Omega_{mag} \ll \Delta$ . This holds if the inverse magnetic correlation length  $m_s$  is small enough because  $\Omega_{mag} \propto m_s$  [32]. In the Fe-based materials, in which  $B_{1g}$  resonance has been observed, the situation is, however, somewhat different: neutron scattering data for  $\text{NaFe}_{1-x}\text{Co}_x\text{As}$  with  $x = 0.045$  show [33] that  $\Omega_{mag} \approx 7$  meV, while  $\Delta = 5 - 5.5$  meV [34], i.e.,  $\Delta$  is somewhat smaller than  $\Omega_{mag}$ . Similarly, for  $\text{Ba}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$  with  $x=0.075$ ,  $\Omega_{mag} \approx 9.5$  meV [35], while  $\Delta = 4.5 - 5$  meV on the electron pocket and 7 meV on the hole pocket [36], so again  $\Delta < \Omega_{mag}$ . In this situation, there is no good reason to treat  $\Gamma_{tr}$  as a constant, independent on  $\Omega_{mag}$ , because two fermionic frequencies in the triangular loop differ by  $\Omega_{mag}$ .

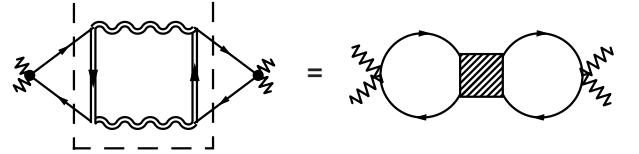


Figure 4. Interpretation of the two-loop AL diagram as consisting of two particle-hole bubbles with zero momentum transfer (unshaded circles), separated by the magnetically-mediated four-fermion interaction. The notations are the same as in Fig. 2. Higher-order terms are shown in Fig. 8.

In view of this complication, we also analyze another scenario for the  $B_{1g}$  resonance below  $T_c$ , namely, that the resonance originates from the  $2\Delta$  singularity of the particle-hole polarization  $\Pi_{B_{1g}}(\Omega)$  in an  $s$ -wave superconductor, like in the pure orbital fluctuation scenario, while spin fluctuations renormalize the original, likely repulsive, coupling  $\lambda$  in Eq. (2) into the effective coupling  $\lambda_{eff}$  (see Figs. 4). This scenario is justified in the opposite limit when  $\Omega_{mag}$  is assumed to be much larger than  $\Delta$ . The strong inequality never holds because  $\Omega_{mag} < 2\Delta$ , but weak inequality may be already enough number-wise. We show that  $\lambda_{eff}$  turns out to be negative (i.e., attractive) near a nematic instability. For negative  $\lambda_{eff}$ ,  $1 + \lambda_{eff}\Pi_{B_{1g}}(\Omega)$  necessary vanishes at some frequency  $\Omega_{res,2}$  below  $2\Delta$ , where  $\text{Im}\Pi_{B_{1g}}(\Omega)$  is zero, and this leads to an excitonic-type resonance in the

$B_{1g}$  Raman response.

Such a scenario has been proposed in earlier works [13 and 23] based on the phenomenological argument that Ising-nematic and orbital order parameters break the same  $C_4$  symmetry and hence are linearly coupled in the Landau functional (see e.g., Ref. [37]). A bilinear term with a constant prefactor  $A$  was argued to give rise to the renormalization of  $\lambda$  into  $\lambda_{eff} = \lambda - A^2 \chi_{I-nem}$ . The latter is obviously negative when  $\chi_{I-nem}$  is large. We compute the renormalization of  $\lambda$  within our microscopic model. We show that  $\lambda_{eff}$  does become negative and scales as  $\chi_{I-nem}$ . However, the prefactor is not  $A^2$  and is nonzero only if one includes the non-analytic dynamical Landau damping term into the spin propagator. If spin-fluctuation propagator is approximated by its static part,  $\lambda_{eff} = 0$ . We explain the difference between the coupling between Pomeranchuk order parameter and two spin fluctuations ( $= A$ ) and between Ising-nematic and orbital (Pomeranchuk) order parameters, which actually gives rise to the renormalization of  $\lambda$ .

The outcome of this study is that the resonance in  $B_{1g}$  Raman intensity is of purely magnetic origin at  $\Omega_{mag} \ll \Delta$ —it is due to the pole in  $\chi_{I-nem}(\Omega)$  at  $\Omega_{res,1}$ . At  $\Omega_{mag} > \Delta$  the resonance of fermionic origin—it develops by the same reason as in purely orbital scenario, due to the emergence of the excitonic pole in the ladder series of particle-hole bubbles with  $B_{1g}$  vertices. However, the attraction between fermions in the  $B_{1g}$  channel comes from magnetically mediated interaction.

In FeSCs,  $\Omega_{mag}$  and  $\Delta$  are comparable, in which case the actual resonance is likely the mixture of the nematic and the excitonic resonances. In  $\text{NaFe}_{1-x}\text{Co}_x\text{As}$  with  $x = 0.045$ , the  $B_{1g}$  peak is seen at 7.1 meV [9], which is below both  $2\Delta$  and  $2\Omega_{mag}$ . Similarly, in  $\text{Ba}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$  with  $x=0.061$ , the  $B_{1g}$  peak is seen at 8.7 meV [12], which is also below both  $2\Delta$  and  $2\Omega_{mag}$ .

The paper is organized as follows: In Sec. II we evaluate analytically the two-loop AL-type diagram for  $B_{1g}$  Raman scattering. In Sec. III we show that in the normal state this contribution to  $\chi_R(\Omega)$  is strongly temperature dependent. The temperature dependence is roughly  $1/T$ . We argue that higher-order terms, which include interactions between pairs of spin fluctuations, replace  $1/T$  dependence into more singular  $1/(T - T_0)$ . In Sec. IV we extend the analysis to the superconducting state. In Sec. IV A we argue that spin excitations with momenta near  $(0, \pi)$  and  $(\pi, 0)$  evolve below  $T_c$  and become magnon-like, with minimal energy  $\Omega_{mag}$ . We compute the two-loop AL diagram assuming that the vertices that couple light to spin fluctuations saturate below  $T_c$ , and show that this contribution to Raman intensity becomes logarithmically singular at  $\Omega = 2\Omega_{mag}$ . We further show that higher-order terms, which include interactions between spin fluctuations, convert logarithmic singularity at  $2\Omega_{mag}$  into a true resonance peak at an energy  $\Omega_{res,1} < 2\Omega_{mag}$ . In Sec. IV B we re-interpret the two-loop AL diagram for  $B_{1g}$  Raman scattering differently, as the contribution from two particle-hole polarization bub-

bles with an effective interaction mediated by spin fluctuations. We compute the magnetically mediated four-fermion interaction  $\lambda_{mag}$  and show that it is attractive. We argue that higher-order terms give rise to an excitonic peak in  $\chi_R(\Omega)$  at  $\Omega_{res,1}$  below  $2\Delta$ . In Sec. IV C we discuss the interplay between this peak and the one coming from fully renormalized nematic susceptibility. In Sec. IV D we present the results of numerical computation of spin-fluctuation contribution to Raman intensity at two-loop and higher orders. In Sec. V we compare AL vertices for the coupling to spin fluctuations in different symmetry channels and show that in the  $B_{2g}$  channel (in the 1-Fe zone) the vertex for the coupling of light to spin fluctuations vanishes by symmetry. The vertex in  $A_{1g}$  channel does not vanish and is of the same order as the vertex in  $B_{1g}$  channel. We show, however, that there is no resonance in  $A_{1g}$  because the interaction between spin fluctuations in this channel is strongly repulsive instead of attractive. We present our conclusions in Sec. VI.

Throughout the paper we will be using band formalism and will be working in the 1-Fe Brillouin zone (BZ).

## II. RAMAN RESPONSE FROM SPIN FLUCTUATIONS

We consider the four-band model of  $\text{NaFe}_{1-x}\text{Co}_x\text{As}$  and  $\text{Ba}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$  with two hole pockets centered at  $k_x = k_y = 0$  and two electron pockets centered at  $(0, \pi)$  and  $(\pi, 0)$  in the 1 Fe Brillouin zone (BZ), see Refs. [38–41]. Excitations near the hole pockets are composed out of  $d_{xz}$  and  $d_{yz}$  orbitals and there is  $90^\circ$  rotation of the orbital content near one Fermi surface compared to the other. Excitations near electron pockets are predominantly composed out of  $d_{xy}$  and  $d_{xz}$  orbitals for the  $(0, \pi)$  pocket and out of  $d_{xy}$  and  $d_{yz}$  orbitals for the  $(\pi, 0)$  pocket [14]. We do not include into consideration the third hole pocket, centered at  $(\pi, \pi)$  in the 1Fe zone, as it is made out of  $C_4$ -symmetric  $d_{xy}$  orbital and does not play any significant role in the analysis of Raman scattering, particularly in  $B_{1g}$  geometry.

The Raman response function can be calculated as a time-ordered average of density operators weighted with Raman form factors:

$$\chi_R(\mathbf{p}, \Omega) = -i \int dt e^{i\Omega t} \langle T \rho_{\mathbf{p}}(t) \rho_{-\mathbf{p}}(0) \rangle, \quad (4)$$

where

$$\rho_{\mathbf{p}} \equiv \sum_{i, \mathbf{k}, \sigma} \gamma_i(\mathbf{k}) c_{i, \mathbf{k}+\mathbf{p}, \sigma}^\dagger c_{i, \mathbf{k}, \sigma}. \quad (5)$$

Here  $i$  represents a band index,  $\sigma$  represents a spin projection of a fermion, and  $\gamma(\mathbf{k})$  is the Raman form factor, which keeps track of the polarizations of the incoming and the outgoing light. The use of light of different polarizations allows the probing of different symmetry channels:  $A_{1g}$ ,  $A_{2g}$ ,  $B_{1g}$ , and  $B_{2g}$ . Note that the  $B_{1g}$  and  $B_{2g}$

channels are interchanged when going from the 1-Fe BZ to the 2-Fe BZ, because the coordinate system is rotated  $45^\circ$  to make the  $k_x$  and  $k_y$  axes coincide with the sides of the square cell. Because the wavelength of light used in the experiments is a few orders of magnitude greater than the lattice constant, the typical values of  $v_F p$  are smaller than typical  $\Omega$ , and it suffices to calculate the susceptibility at  $p \rightarrow 0$ , i.e., compute  $\chi_R(\Omega) \equiv \chi_R(p \rightarrow 0, \Omega)$ .

Without the final state interaction, the Raman response involving a pair of spin fluctuations with momenta near  $(0, \pi)$  and  $(\pi, 0)$  (the difference between the centers of electron and hole pockets) is given by the diagram shown in Fig. 2. Since light can couple to each hole and electron band, there are several diagrams of this kind with two fermionic lines from one of hole pockets or from one of electron pockets (see Fig. 16 below). The combined contribution from these diagrams takes the form

$$\chi_R(\Omega) = -i \int \frac{d^2 \mathbf{q} d\nu}{(2\pi)^3} \Gamma_{tr,l}^2(\mathbf{q}, \nu) f_l \chi^s(\mathbf{Q}_l + \mathbf{q}, \nu) \times \chi^s(\mathbf{Q}_l + \mathbf{q}, \nu + \Omega), \quad (6)$$

where  $\Gamma_{tr,l}$  defines the vertex for the coupling between light and spin fluctuations (see Fig. 5),  $\chi^s$  is the propagator of spin fluctuations,  $l = 1, 2$  with  $\mathbf{Q}_1 = (\pi, 0)$ ,  $\mathbf{Q}_2 = (0, \pi)$ , and  $f_l$  is the symmetry factor, e.g.,  $f_l = \sigma_{ll}^z$  for  $B_{1g}$  geometry.

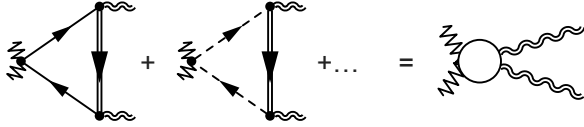


Figure 5. The AL vertex  $\Gamma_{tr}$  for the coupling of light to spin fluctuations. Single solid and dashed lines represent excitations from one of the two hole bands and the double solid line represents excitations from one of the two electron bands. When typical internal frequencies in the triangle made out of fermionic Green's functions are larger than typical frequencies of spin fluctuations,  $\Gamma_{tr}$  can be approximated by a constant (a circle on the right-hand side of the figure).

The vertex  $\Gamma_{tr,l}(\mathbf{q}, \nu)$  is composed out of three fermionic Green's functions, the Raman factor  $\gamma(\mathbf{k})$ , and two vertices for the coupling between fermions and spin-fluctuations:  $g_{sf} \gamma_{ij}(\mathbf{k}, \mathbf{k} + \mathbf{Q}_l - \mathbf{q})$ , where  $g_{sf}$  is the spin-fermion coupling (or order of the intra-orbital Hubbard  $U$ ) and  $\gamma_{ij}(\mathbf{k}, \mathbf{k} + \mathbf{Q}_l - \mathbf{q})$  are coherence factors associated with the transformation from orbital to band basis for fermions from bands  $i$  and  $j$  ( $i, j = 1, 2$ ). In a superconducting state one must include diagrams containing anomalous Green's functions, which create or destroy particles in pairs, and sum over all allowed combinations of normal and anomalous functions. For any particular combination of normal and anomalous functions, the tri-

angular vertex takes the form

$$\Gamma_{tr,l}(\mathbf{q}, \nu) = g_{sf}^2 \int \frac{d^2 \mathbf{k} d\omega}{(2\pi)^3} \gamma_i(\mathbf{k}) \gamma_{ij}^2(\mathbf{k} + \mathbf{Q}_l - \mathbf{q}) \tilde{G}_i(\mathbf{k}, \omega) \tilde{G}_i(\mathbf{k}, \omega + \Omega) \tilde{G}_j(\mathbf{k} + \mathbf{Q}_l - \mathbf{q}, \omega - \nu), \quad (7)$$

where  $\tilde{G}_{i,j}$  are either normal ( $G$ ) or anomalous ( $F$ ) Green's functions with  $i$  from a hole pocket and  $j$  from an electron pocket or vice versa. The index of the electron band is equal to  $l$ .

### III. THE VERTEX FUNCTION IN THE NORMAL STATE AND THE TEMPERATURE DEPENDENCE OF THE $B_{1g}$ SUSCEPTIBILITY ABOVE $T_c$

In this section we compute the Raman vertex function in the normal state and obtain the temperature dependence of the real part of the Raman susceptibility in the static limit. Here and in Sec. IV we focus on  $B_{1g}$  scattering geometry with  $\gamma(\mathbf{k}) \propto \cos \mathbf{k}_x - \cos \mathbf{k}_y$  and do not explicitly write symmetry factors in the Raman vertex and in the vertices relating fermions with spin fluctuations. We will discuss these symmetry factors and different geometries in Sec. V. We also neglect for simplicity the eccentricity of the electron pockets and set all Fermi surfaces to be circles of the same size. This approximation simplifies calculations, but does not qualitatively affect the temperature dependence compared to a generic case in which the pockets are different. In the static limit Eq. (7) in the normal state reduces to

$$\Gamma_{tr}(\mathbf{q}, \nu_n) = -AT \sum_{\omega_m} \int \frac{d^2 \mathbf{k}}{(2\pi)^2} \frac{1}{(i\omega_m - \xi_{\mathbf{k}}^i)^2} \frac{1}{i(\omega_m - \nu_n) - \xi_{\mathbf{k}-\mathbf{q}}^j}, \quad (8)$$

where  $A \sim g_{sf}^2$ . For concreteness we assume that  $i$  is a hole band and  $j$  is an electron band. The hole and electron dispersions are given by  $\xi_{\mathbf{k}}^i = \mu - \frac{k^2}{2m} = -\xi_{\mathbf{k}}^j$ , where  $\mu$  is the chemical potential. At  $\mathbf{q} = 0$  and  $\nu_m = 0$ ,  $\Gamma_{tr}$  is given by

$$\Gamma_{tr} = \frac{Am}{16\pi T} f\left(\frac{\mu}{2T}\right).$$

The scaling function  $f(x) = \tanh(x)/x$  is close to 1 for  $x < 1$ , i.e., for  $T > \mu/2$ . In this temperature range  $\Gamma_{tr} \approx Am/(16\pi T)$  scales as  $1/T$ . This has been noticed before [42]. At larger  $x$  (smaller  $T$ ),  $\Gamma_{tr}$  tends to a constant [31]. At a nonzero  $\mathbf{q}$  and  $\nu_m$  the expression for  $\Gamma_{tr}$  becomes more complex, but as long as  $\nu_m = O(T)$  and  $|\mathbf{q}| \leq k_F$ , the functional form remains the same.

We next compute the convolution of two spin fluctuations in the normal state. There is no controllable way to obtain the spin-fluctuation propagator starting from the fermion-fermion interaction. The RPA procedure is

often used, but it selects particular series of ladder and bubble diagrams in the particle-hole channel and neglects contributions from the particle-particle channel. The latter are, however, not small, even at perfect nesting [43]. Besides, in a general case of hole and electron pockets of different sizes and geometry, the static propagator of spin fluctuations comes from fermions with energies of order bandwidth, for which the low-energy expansion is not applicable. In view of this complication, we adopt the same approach as in earlier works on the spin-fermion model [44] and assume phenomenologically that the static part of the spin-fluctuation propagator has a regular Ornstein-Zernike form  $\chi_{ij}^s(\mathbf{q} + \mathbf{Q}, 0) = 1/(q^2 + m_s^2)$ , where  $m_s$  is the inverse magnetic correlation length (the overall factor in  $\chi^s$  is incorporated into the spin-fermion coupling). The dynamical part of  $\chi^s$ , however, comes from low-energy fermions and can be obtained by computing the dynamical part of particle-hole polarization bubble made of fermions near a hole and an electron pocket, separated by  $\mathbf{Q}$ . Then

$$\chi^s(\mathbf{Q} + \mathbf{q}, \nu_m) = \frac{1}{m_s^2 + q^2 + \gamma \Pi_Q(\nu_m)}, \quad (9)$$

where  $\gamma = mg_{sf}/(2\pi)$ ,  $g_{sf}$  is the spin-fermion coupling [44], and  $\Pi_Q(\nu_m) = \Pi(\mathbf{Q}, \nu_m) - \Pi(\mathbf{Q}, 0)$ , where  $\Pi(\mathbf{Q}, \nu_m)$  is the dynamical polarization bubble at momentum transfer  $\mathbf{Q}$ . The polarization bubble  $\Pi_Q(\nu_m)$  is logarithmic in  $\nu_m$  because it is the convolution of fermions from hole and electron bands with opposite sign of the dispersion [46]. We computed  $\Pi_Q(\nu_m)$  numerically and found that it can be well approximated by

$$\Pi_Q(\nu_m) = \log \left( 3.57 \frac{|\nu_m|}{2\pi T} \right) \quad (10)$$

starting already from the lowest nonzero Matsubara frequency. Substituting this into (9) and evaluating the convolution of the two dynamical spin susceptibilities with the same momentum and frequency we obtain

$$T \sum_{\nu_m} \int \frac{d^2 \mathbf{q}}{(2\pi)^2} \chi^s(\mathbf{q} + \mathbf{Q}, \nu_m) \chi^s(\mathbf{q} + \mathbf{Q}, \nu_m) \\ \propto T(m_s)^{-2} \left( 1 + \sum_{m \neq 0} \frac{1}{1 + \gamma(m_s)^{-2} \log(3.57|m|)} \right) \quad (11)$$

The coupling constant  $\gamma$  cannot be calculated within the theory, but is generally of order 1. Assuming that this is the case, we find that the dominant contribution to the sum over bosonic Matsubara frequencies comes from the term with  $\nu_n = 0$ , at least when the inverse magnetic correlation length  $m_s < 1$ . The convolution of the two  $\chi^s$  then gives, up to a constant prefactor,  $T/m_s^2$  [21]. Combining this with  $\Gamma_{tr}^2 \propto 1/T$ , we obtain that the contribution to the static  $B_{1g}$  Raman susceptibility from the processes involving a pair of spin fluctuations with momenta near  $(0, \pi)$  and  $(\pi, 0)$  and two triple vertices (i.e.,

from the diagram in Fig. 2) is given by

$$\chi_R(\Omega = 0) \propto \frac{1}{T m_s^2} \quad (12)$$

Outside the  $T$  range near a magnetic transition, the temperature dependence of  $m_s$  is weak, and  $\chi_R(\Omega = 0)$  scales roughly as  $1/T$ .

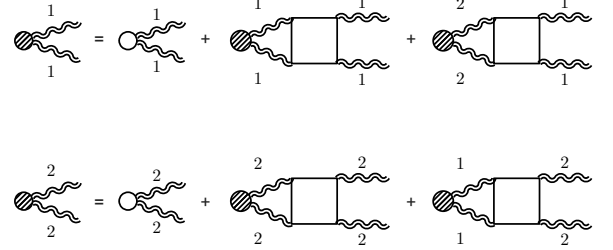


Figure 6. Ladder series of renormalizations of the interaction of light with spin fluctuations with momenta near  $\mathbf{Q}_1 = (\pi, 0)$  (labeled as 1) and near  $\mathbf{Q}_2 = (0, \pi)$  (labeled as 2). The interaction vertices are made out of fermions from different bands (see Fig. 7).

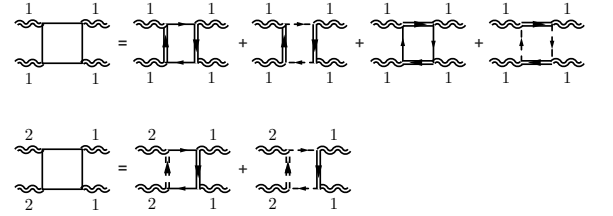


Figure 7. The structure of the vertices for the interaction between spin fluctuations with near  $\mathbf{Q}_1 = (\pi, 0)$  (labeled as 1) and near  $\mathbf{Q}_2 = (0, \pi)$  (labeled as 2). The single solid and dashed lines represent excitations from one of the two hole bands and the double solid and dashed lines represent excitations from one of the two electron bands.

#### A. Higher-order contributions to Raman susceptibility

We next consider how Eq. (12) changes once we include the interactions between pairs of spin fluctuations. These interactions either leave bosonic momenta near a particular  $\mathbf{Q}_1 = (\pi, 0)$  or  $\mathbf{Q}_2 = (0, \pi)$ , or transfer both momenta from  $\mathbf{Q}_1$  to  $\mathbf{Q}_2$  or vice versa. In the latter case, the process belongs to the umklapp category and is allowed because  $2\mathbf{Q}_1 = 2\mathbf{Q}_2$  up to reciprocal lattice vector. We show the corresponding diagrams in Fig. 6. Each interaction vertex is given by the convolution of four fermionic propagators (see Fig. 7). These vertices have been computed in Ref. [28] in the limit when bosonic frequencies are set to zero. We argued above that this approximation is justified for the two-loop diagram, and

we now assume that this also holds for higher-order processes. We then borrow the results from Ref. [28], which demonstrated that the effective interaction in the  $B_{1g}$  channel is a negative (attractive)  $2g$ , which scales with temperature as  $1/T^4$ .

We show explicitly how to solve the coupled set of ladder equations for the fully renormalized AL vertices with momenta  $\mathbf{Q}_1$  and  $\mathbf{Q}_2$  later in Sec. V (where we compare  $A_{1g}$  and  $B_{1g}$  channels), because the set involves both  $A_{1g}$  and  $B_{1g}$  components. Here we just present the result: at small  $\Omega$ , interactions between spin-fluctuations change the two-loop AL Raman vertex in  $B_{1g}$  geometry into

$$\chi_R(\Omega) = \Gamma_{tr}^2 \frac{\chi^2(\Omega)}{1 + 2g\chi^2(\Omega)} = \Gamma_{tr}^2 \chi_{I-nem}(\Omega), \quad (13)$$

where  $\chi^2(\Omega)$  is the short notation for the convolution of two spin-fluctuation propagators with relative frequency  $\Omega$  and we defined

$$\chi_{I-nem}(\Omega) = \frac{\chi^2(\Omega)}{1 + 2g\chi^2(\Omega)}. \quad (14)$$

Using  $g \sim 1/T^4$ ,  $\chi^2 \propto T$ ,  $\Gamma_{tr} \propto 1/T$ , and introducing  $T_0$  as a temperature at which  $2g\chi^2 = 1$ , we obtain

$$\chi_R(0, T) \sim \frac{(m_s)^{-2}}{T - T_0} \left( \frac{T^2}{T^2 + TT_0 + T_0^2} \right), \quad (15)$$

which for  $T > T_0$  is rather well approximated by

$$\chi_R \sim \frac{1}{T - T_0}. \quad (16)$$

In Sec. V we show that singular  $1/(T - T_0)$  dependence only holds for the  $B_{1g}$  Raman vertex. In other channels, Raman intensity from the coupling to spin fluctuations either vanishes by symmetry or is substantially reduced by interaction between spin fluctuations.

The  $1/(T - T_0)$  behavior of the  $B_{1g}$  Raman vertex is quite consistent with the experimental observations for  $\text{NaFe}_{1-x}\text{Co}_x\text{As}$ ,  $\text{EuFe}_2\text{As}_2$ , and  $\text{Ba}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$  [9, 10, 12, and 13]. In these materials  $T_0$  is positive at doping below a certain  $x_0$ . This  $T_0$  would be the temperature of Ising-nematic instability in the absence of (i) superconductivity and (ii) coupling to phonons. Superconductivity obviously cuts  $1/(T - T_0)$  behavior at  $T_c$ , if  $T_c$  is larger than  $T_0$ . The coupling to static phonons shifts the the coupling  $g$  to  $g_{eff} = g + (\Gamma_{tr}\lambda_{ph}/C_{ph})^2/2$  [see Eq. (1)] and hence shifts the temperature of the Ising-nematic instability to  $T_{nem} > T_0$ . We remind that this shift is not present in the "static"  $\chi_R$ , extracted from the measured  $\text{Im } \chi_R(\Omega)$  by Kramers-Kronig transformation, because the contribution from phonons rapidly drops at nonzero  $\Omega$  and is negligibly small for frequencies at which  $\text{Im } \chi_R(\Omega)$  has been measured. As a result, "static"  $B_{1g}$  susceptibility extracted from the Raman data increases upon decreasing  $T$  but does not diverge even at  $T_{nem}$ .

Note also that  $T_0$  is positive as long as magnetic order is a stripe, otherwise  $g > 0$  and the final state interaction reduces rather than enhances the Raman intensity.

Recent studies have shown [45] that  $g$  does change sign as doping increases, hence one should expect that  $T_0$  will change sign from positive to negative above a certain doping. This is fully consistent with the data [9].

#### IV. THE RESONANCE IN $B_{1g}$ CHANNEL BELOW $T_c$

We now turn to the superconducting state. We first argue in Sec. IV A that under certain conditions the resonance in  $B_{1g}$  Raman response is due to the development of the pole in  $\chi_{I-nem}(\Omega)$ , given by (14). In Sec. IV B we consider another scenario for the resonance. Namely, we re-interpret the two-loop diagram as containing two dynamical particle-hole polarization bubbles with zero momentum transfer,  $\Pi_{B_{1g}}(\Omega)$ , coupled by an effective interaction mediated by spin fluctuations (see Figs. 4, 8). This effective interaction renormalizes  $\lambda$  in Eq. 2 into  $\lambda_{eff}$ , and we show that  $\lambda_{eff}$  becomes negative. As we said in the Introduction, for negative coupling the system develops an excitonic resonance in the superconducting state, at a frequency below twice the superconducting gap.

##### A. Resonance due to the pole in $\chi_{I-nem}$

To analyze the form of  $\chi_{I-nem}$  at  $T = 0$  in a superconductor we need to know the form of  $\chi^s(\mathbf{Q} + \mathbf{q}, \nu)$  along the real frequency axis. As in earlier works, we assume that the symmetry of the superconducting order parameter is  $s^{+-}$ . Superconductivity does not affect the static form of  $\chi^s$  as it generally comes from high-energy fermions, but changes the form of the dynamical term  $\Pi_Q(\nu) = \Pi_s(\mathbf{Q}, \nu) - \Pi_s(\mathbf{Q}, 0)$  in Eq. (9) (converted to real frequencies), as this term now contains the sum of  $G_s G_s$  and  $F_s F_s$  terms. Approximating fermionic dispersion in the same way as before we obtain

$$\Pi_s(\mathbf{Q}, \nu) = \frac{1}{2} \int \frac{d^2\mathbf{k}}{(2\pi)^2} \left[ \frac{1}{\nu + 2E_{\mathbf{k}} - i\eta} - \frac{1}{\nu - 2E_{\mathbf{k}} + i\eta} \right], \quad (17)$$

where  $E_{\mathbf{k}}^2 = \xi_{\mathbf{k}}^2 + \Delta^2$  and  $\xi_{\mathbf{k}} = \mu - \frac{k^2}{2m}$  (note that we define  $\Pi$  without a spin factor of 2). In principle, in evaluating  $\Pi_s(\mathbf{Q}, \nu)$  one has to include also  $G_s F_s$  terms and combine renormalizations in the particle-hole and the particle-particle channels because in the superconducting state particles and holes are mixed [47]. Previous work on the subject [48], however, has shown that as long as all the interactions are repulsive, the effect of inclusion of these extra terms is minimal in the case of Fe pnictides and merely shifts the resonance frequency (see below) down by a few percentage points.

The straightforward analysis shows that  $\text{Im } \Pi_Q(\nu)$  vanishes for  $|\nu| < 2\Delta$  because the excitations are gapped. At

$|\nu| = 2\Delta$ ,  $\text{Im } \Pi_Q(\nu)$  undergoes a discontinuous jump to a finite value and  $\text{Re } \Pi_Q(\mu)$  diverges logarithmically. The divergence of the real part of  $\Pi_Q(\nu)$  at  $\nu = 2\Delta$  implies that the denominator in (9) must vanish at some frequency below  $2\Delta$ , thus creating a pole in  $\chi^s(\mathbf{Q} + \mathbf{q}, \nu)$ . Specifically, for a given  $\mathbf{q}$ ,  $\text{Im } \chi^s(\mathbf{Q} + \mathbf{q}, \nu)$  has sharp peak at frequency  $\nu_{res}(\mathbf{q})$  and  $\text{Re } \chi^s(\mathbf{Q} + \mathbf{q}, \nu)$  diverges. This is indeed nothing but the spin resonance in an  $s^{+-}$  superconductor [49]. Because time-ordered  $\Pi_Q(\nu)$  is an even function of  $\nu$ , it follows that for a given  $\mathbf{q}$ , time-ordered  $\chi^s(\mathbf{Q} + \mathbf{q}, \nu)$  has two simple poles at  $\nu = \pm \nu_{res}(\mathbf{q})$ . Then  $\chi^s$  can be written as

$$\chi^s(\mathbf{Q} + \mathbf{q}, \nu) = \frac{a(\mathbf{q}, \nu)}{[\nu + \nu_{res}(\mathbf{q})][\nu - \nu_{res}(\mathbf{q})]}, \quad (18)$$

where  $a(\mathbf{q}, \nu)$  is some analytic function, which is also even in  $\nu$ .

We now turn to the Raman susceptibility from the two-loop diagram, Eq. (6). We assume and then verify that the triangular Raman vertex  $\Gamma_{tr}$  can be approximated by a constant and taken out of the integral for  $\chi_R$ . The  $1/T$  temperature dependence of  $\Gamma_{tr}$  is obviously cut by  $T_c$ , i.e., it remains finite at  $T = 0$ . Whether it can be taken out of the integral over the bosonic frequency is a more subtle issue and we discuss it at the end of this section.

With  $\Gamma_{tr}$  approximated by a constant, the expression for the Raman susceptibility takes the form

$$\chi_R(\Omega) = \Gamma_{tr}^2 \chi^2(\Omega), \quad (19)$$

where

$$\chi^2(\Omega) = -i \int \frac{d^2 \mathbf{q} d\nu}{(2\pi)^3} \sigma_{ii}^z \chi^s(\mathbf{Q}_i + \mathbf{q}, \nu) \chi^s(\mathbf{Q}_i + \mathbf{q}, \nu + \Omega). \quad (20)$$

where we remind that  $i = 1, 2$ ,  $\mathbf{Q}_1 = (\pi, 0)$ ,  $\mathbf{Q}_2 = (0, \pi)$ , and  $\sigma_{ii}^z$  is present because we consider  $B_{1g}$  geometry.

Substituting  $\chi_s$  from (18) into Eq. (20) and evaluating the frequency integral, we obtain, neglecting symmetry factors,

$$\begin{aligned} \chi^2(\Omega) = & - \int \frac{d^2 \mathbf{q}}{(2\pi)^2} \frac{a[\mathbf{q}, \nu_{res}(\mathbf{q})]}{2\nu_{res}(\mathbf{q})\Omega} \\ & \times \left[ \frac{a[\mathbf{q}, \Omega - \nu_{res}(\mathbf{q})]}{\Omega - 2\nu_{res}(\mathbf{q})} + \frac{a[\mathbf{q}, \Omega + \nu_{res}(\mathbf{q})]}{\Omega + 2\nu_{res}(\mathbf{q})} \right]. \end{aligned} \quad (21)$$

This formula shows that for each momentum  $\mathbf{q}$  there is an enhancement of the response at twice the frequency  $\nu_{res}(\mathbf{q})$ . We define the minimum value of  $\nu_{res}(\mathbf{q})$  as  $\Omega_{mag}$ . A simple experimentation with the momentum integral shows that  $\text{Im } \chi^2$  is small at  $\Omega < 2\Omega_{mag}$ , but enhances sharply at  $\Omega \geq 2\Omega_{mag}$ . In order to illustrate this effect more concretely, we adopt a simple model for the dispersion of the pole. Namely, we set  $\nu_{res}(\mathbf{q}) = \Omega_{mag} + \alpha q^2$ . Integrating in Eq. (22) over  $q$  and substituting the result into (19) we obtain

$$\chi_R(\Omega) = \frac{\Gamma_{tr}^2 a^2}{8\pi\alpha\Omega^2} \log \frac{4\Omega_{mag}^2}{4\Omega_{mag}^2 - \Omega^2}. \quad (23)$$

We see that in the two-loop approximation,  $\text{Im } \chi_R(\Omega)$  undergoes a jump from zero to a finite value at  $\Omega = 2\Omega_{mag}$ . The real part of the Raman susceptibility  $\text{Re } \chi_R(\Omega)$  diverges logarithmically at this frequency. Below we verify this result by evaluating (20) numerically.

We next follow the same path as in the normal state and include higher-order diagrams (Fig. 3) with the interactions between the two spin fluctuations, i.e., replace the two-loop result  $\chi_R = \Gamma_{tr}^2 \chi^2(\Omega)$  by

$$\chi_R(\Omega) = \Gamma_{tr}^2 \chi_{I-nem}(\Omega) = \Gamma_{tr}^2 \frac{\chi^2(\Omega)}{1 + 2g\chi^2(\Omega)}. \quad (24)$$

Because  $\text{Re } \chi^2$  diverges upon approaching  $2\Omega_{mag}$  from below,  $\text{Im } \chi^2$  vanishes below  $2\Omega_{mag}$ , and  $g < 0$ , the full  $\text{Im } \chi_R(\Omega)$  has a true pole at some frequency  $\Omega = \Omega_{res,1} < 2\Omega_{mag}$ .

We now verify the approximation that  $\Gamma_{tr}$  can be taken out of the integral over the bosonic frequency  $\nu$ . The triangular vertex contains one internal frequency  $\omega$  and two external ones:  $\Omega$ , at which we probe the Raman response, and the bosonic frequency  $\mu$ . For  $\Omega$  we take the resonance frequency  $\Omega_{res,1} < 2\Omega_{mag}$ . Typical bosonic frequency  $\nu \sim \Omega_{mag}$  and typical  $\omega$  is  $\Delta$ . Obviously then, the AL vertex  $\Gamma_{tr}$  is independent on  $\nu$  if  $\Omega_{mag}$  is much smaller than  $\Delta$ , i.e., when internal energy in the AL diagram made out of three fermionic Green's functions is much larger than both external frequencies.

The condition  $\Omega_{mag} \ll \Delta$  is satisfied when the inverse magnetic correlation length  $m_s$  is small enough because  $\Omega_{mag} \propto m_s$  [32]. As we said in the Introduction, in Fe-based materials, where  $B_{1g}$  resonance has been observed, the situation is somewhat different: neutron scattering data for  $\text{NaFe}_{1-x}\text{Co}_x\text{As}$  with  $x = 0.045$  show [33] that  $\Omega_{mag} \approx 7$  meV, while  $\Delta = 5 - 5.5$  meV [34], i.e.  $\Delta$  is somewhat smaller. In this situation, there is no good reason to treat  $\Gamma_{tr}$  as a constant.

In the next section we analyze another scenario, which is justified in the opposite limit when  $\Omega_{mag}$  is much larger than  $\Delta$ .

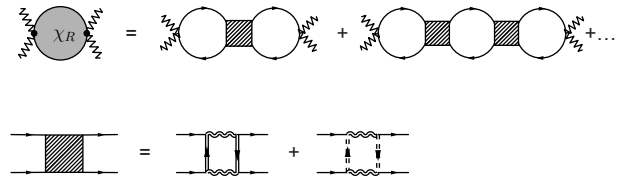


Figure 8. Reinterpretation of series of AL diagrams for  $\chi_R$  with interactions between spin fluctuations as series consisting of multiple particle-hole bubbles with zero momentum and finite frequency transfer (unshaded circles), separated by effective interactions mediated by spin-fluctuations. Each interaction vertex (shaded rectangle) is the convolution of two fermionic and two bosonic propagators. A particular subset of diagrams is shown, with fermions from one of the hole band (solid lines). Double solid and double dashed lines describe fermions from two electron bands.



### B. Another interpretation of AL contribution to the $B_{1g}$ Raman response below $T_c$ .

In this section we look at the second-order AL diagram from Fig. 2 through different lenses. Namely, we abandon the approximation in which  $\Gamma_{tr}$  is treated as a constant and instead use the fact that on both ends incoming and outgoing fermionic momenta are identical (either  $\mathbf{k}$  or  $\mathbf{p}$ ), while frequencies differ by  $\Omega$ , and re-interpret this diagram as consisting of the product of two particle-hole polarization operators at zero transferred momentum and finite frequency,  $\Pi_{B_{1g}}(\Omega)$  [the same as in Eq. (2)], separated by magnetically mediated effective interaction  $\lambda_{mag}$  (see Fig. 8). The latter is the convolution of two fermionic and two bosonic propagators. Viewed this way, the two-loop AL diagram has the same structure as the two-loop diagram from RPA series in Eq. (2). Accordingly,  $\lambda_{mag}$  and the bare  $\lambda$  are combined into  $\lambda_{eff} = \lambda + \lambda_{mag}$ . If the combined  $\lambda_{eff}$  is negative,  $1 + \lambda_{eff}\Pi_{B_{1g}}(\Omega)$  necessary vanishes at some frequency below  $2\Delta$  because  $\text{Im } \Pi_{B_{1g}}(\Omega)$  vanishes below  $2\Delta$  and  $\text{Re } \Pi_{B_{1g}}(\Omega)$  diverges as  $1/\sqrt{4\Delta^2 - \Omega^2}$  when  $|\Omega|$  approaches  $2\Delta$  from below. The vanishing of  $1 + \lambda_{eff}\Pi_{B_{1g}}(\Omega)$  implies that Raman intensity has an excitonic resonance at  $\Omega = \Omega_{res,2}$ .

The representation of the two-loop AL diagram from Fig. 2 as  $\lambda_{mag}\Pi_{B_{1g}}^2(\Omega)$  with a constant  $\lambda_{mag}$  is again an approximation because the result for the convolution of two fermionic and two bosonic propagators generally depends on external momentum in frequency. The singular behavior of the particle-hole polarization bubble  $\Pi_{B_{1g}}(\Omega)$  at  $\Omega \approx 2\Delta$  comes from internal frequencies near  $\pm\Delta$ . Internal frequencies in the fermionic-bosonic loop for  $\lambda_{mag}$  are of order  $\Omega_{mag}$ . If  $\Omega_{mag}$  is much larger than  $\Delta$ , a typical internal frequency is much larger than a typical external frequency. The latter can then be sent to zero, in which case  $\lambda_{mag}$  becomes a constant. The frequency  $\Omega_{mag} < 2\Delta$  and hence it can be at most twice  $\Delta$ . But number-wise this may be sufficient to treat  $\lambda_{mag}$  as a constant. The same distinction holds for internal/external momenta, and the result is that, to the same accuracy,  $\lambda_{eff}$  can be evaluated by placing external momenta on the Fermi surface.

We compute  $\lambda_{eff}$  first in the normal state and then in a superconductor, assuming formally that  $\Omega_{mag} \gg \Delta$ . To simplify calculations, we set  $\mu = 0$ , i.e., assume that the size of hole/electron pockets is infinitesimally small. The argument is that, if  $\lambda_{eff}$  has a definite sign in this limit, then, by continuity, the sign should remain the same at a small but finite  $\mu$ .

In the normal state, the coupling  $\lambda_{mag}$  is given by

$$\lambda_{mag} = -g_{sf}^2 \int \frac{d^2\mathbf{q}d\nu}{(2\pi)^3} G^2(\mathbf{q}, \nu) (\chi^s(\mathbf{q}, \nu))^2, \quad (25)$$

where  $\nu$  is the Matsubara frequency. For definiteness, we take fermions from one of the electron bands, i.e., use  $G(\mathbf{q}, \nu) = 1/(i\nu - q^2/(2m))$ . We verified that  $\lambda_{mag}$

does not change if we instead take fermions from the hole band. For the dynamical spin susceptibility we use Landau-overdamped form extended to Matsubara frequencies:  $\chi(\mathbf{q}, \nu) = 1/(q^2 + m_s^2 + |\nu|/\nu_0)$ , where  $\nu_0$  is a positive constant.

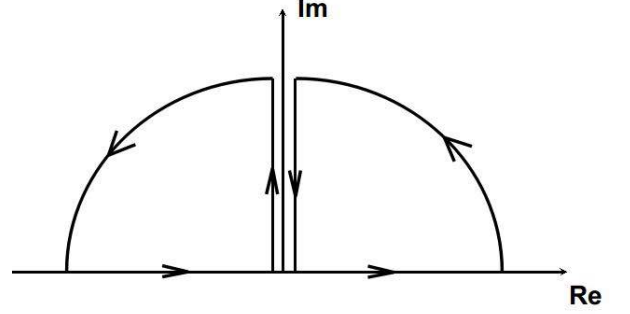


Figure 9. Contour of integration over frequency  $\nu$  in Eq. 26.

Substituting the forms of bosonic and fermionic propagators into (25) we obtain

$$\begin{aligned} \lambda_{mag} &= -g_{sf}^2 \int \frac{d^2\mathbf{q}d\nu}{(2\pi)^3} \frac{1}{[i\nu - \mathbf{q}^2/(2m)]^2} \frac{1}{(\mathbf{q}^2 + m_s^2 + |\nu|/\nu_0)^2} \\ &= \frac{\nu_0^2 m}{(2\pi)^2} \int_0^\infty dx \int_{-\infty}^\infty d\nu \frac{1}{(\nu + ix)^2} \frac{1}{(|\nu| + \nu_0 m_s^2 + \nu_0 mx)^2}. \end{aligned} \quad (26)$$

The double pole in the fermionic Green's function is located at  $\nu = -ix$ . It is then convenient to evaluate the frequency integral by closing the integration contour over the upper half plane of complex  $\nu$  (see Fig. 9). The integrand vanishes at  $|\nu| \rightarrow \infty$ , and if  $\chi^s$  was an analytic function of  $\nu$ ,  $\lambda_{mag}$  would be zero. But this is not the case because the Landau damping term contains  $|\nu| = \sqrt{\nu^2}$ , which is non-analytic function of  $\nu$  along imaginary axis in both half planes. Choosing the integration contour as shown in Fig. 9 and using  $|iz + \epsilon| = iz \text{sgn } \epsilon$ , we obtain after a simple algebra

$$\begin{aligned} \lambda_{mag} &= -ig_{sf}^2 \frac{\nu_0^2 m}{(2\pi)^2} \int_0^\infty dx \int_0^\infty dz \frac{1}{(z+x)^2} \\ &\times \left[ \frac{1}{(iz + \nu_0 m_s^2 + \nu_0 mx)^2} - \frac{1}{(-iz + \nu_0 m_s^2 + \nu_0 mx)^2} \right] \\ &= -\frac{4g_{sf}^2 \nu_0^2 m}{(2\pi)^2} \\ &\times \int_0^\infty dx \int_0^\infty dz \frac{z(\nu_0 m_s^2 + \nu_0 mx)}{(z+x)^2 [z^2 + \nu_0^2 (m_s^2 + mx)^2]^2}. \end{aligned} \quad (27)$$

The integrand is positive, hence  $\lambda_{mag} < 0$ . Estimating the integral, we obtain  $\lambda_{mag} \propto 1/m_s^4$ , i.e.,  $\lambda_{mag}$  strongly increases near the magnetic instability.

In the superconducting state, we represent spin-fluctuation propagator by Eq. (18), i.e., by  $\chi(\mathbf{q}, \nu) = a/[\nu^2 + \nu_{res}^2(\mathbf{q})]$ , and use  $\nu_{res}(\mathbf{q}) = \Omega_{mag} + \alpha q^2$ . We assume and then verify that typical  $\nu$  in the integral for

$\lambda_{mag}$  are of order  $\Omega_{mag}$ . Because we assume  $\Omega_{mag} \gg \Delta$ , we can still use normal state Green's functions for fermions. Substituting into (25) we obtain

$$\begin{aligned} \lambda_{mag} &= -g_{sf}^2 \int \frac{d^2 \mathbf{q} d\nu}{(2\pi)^3} \frac{1}{[i\nu - \mathbf{q}^2/(2m)]^2} \frac{a}{[\nu^2 + (\Omega_{mag} + \alpha q^2)^2]^2} \\ &= \frac{g_{sf}^2}{4\pi^2} \frac{ma}{\Omega_{mag}^4} \int_0^\infty dx \int_{-\infty}^\infty d\nu \frac{1}{(\nu + ix)^2} \frac{1}{[\nu^2 + (1 + \beta x)^2]^2}. \end{aligned} \quad (28)$$

where  $\beta = 2m\alpha$  is a dimensionless parameter. The integrand, viewed as a function of  $\nu$ , contains two double poles in the lower-half plane, at  $\nu = -ix$  and at  $\nu = -i(1 + \beta x)$ , and the double pole in the upper half-plane, at  $\nu = i(1 + \beta x)$ . The last two double poles come from  $\chi^2$ . Evaluating the frequency integral by standard means, we obtain after simple algebra that

$$\lambda_{mag} = -\frac{g_{sf}^2}{8\pi} \frac{ma}{\Omega_{mag}^4} \int_0^\infty dx \frac{3 + x(1 + 3\beta)}{(1 + \beta x)^3 [1 + (1 + \beta)x]^3}. \quad (29)$$

The integrand is positive for all  $x > 0$ , hence  $\lambda_{mag} < 0$ , like in the normal state. Furthermore, because  $\Omega_{mag} \propto m_s$ , we still have  $\lambda_{mag} \propto 1/m_s^4$ .

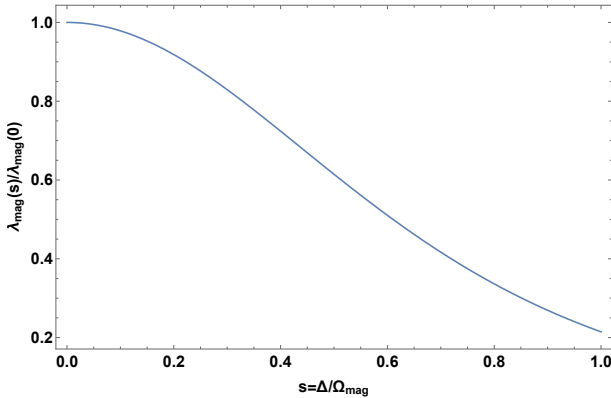


Figure 10. The dependence of the effective spin-mediated four-fermion interaction  $\lambda_{mag}(s)$  on  $s = \Delta/\Omega_{mag}$ , see Eq. (30). We set  $\beta = 1$  for definiteness.

Like we said, the condition  $\Omega_{mag} \gg \Delta$  is not realized because  $\Omega_{mag} \leq 2\Delta$ . To estimate how  $\lambda_{mag}$  changes when  $\Omega_{mag}$  and  $\Delta$  become comparable, we use the fact that  $2\Delta$  singularity in the particle-hole bubble comes from fermions with frequencies near  $\pm\Delta$  and evaluate  $\lambda_{mag}$  in the superconducting state for the case when the frequencies of the two bosonic propagators differ by  $\Omega_m = 2\Delta$ . The calculation is straightforward, and the result is that  $\lambda_{mag}$  becomes a function of  $s = \Delta/\Omega_{mag}$ .

The dependence on  $s$  is given by

$$\begin{aligned} \lambda_{mag}(s) &= -\frac{g_{sf}^2}{8\pi} \frac{ma}{\Omega_{mag}^4} \\ &\times \int_0^\infty dx \frac{[(1 + (1 + \beta)x)[3 + x(1 + 3\beta)] - s^2]}{(1 + \beta x)[(1 + \beta x)^2 + s^2]\{[1 + (1 + \beta)x]^2 + s^2\}^2} \end{aligned} \quad (30)$$

We plot  $\lambda_{mag}(s)$  in Fig. 10. We see that  $\lambda_{mag}(s)$  drops when  $s = \Delta/\Omega_{mag}$  increases, but the ratio  $\lambda_{mag}(s)/\lambda_{mag}(0)$  remains of order one when  $\Delta$  and  $\Omega_{mag}$  become comparable.

Combining this last observation with the fact that  $\lambda_{mag}(0) \propto 1/m_s^4$ , we conclude that for small enough  $m_s$ ,  $|\lambda_{mag}|$  is definitely larger than the bare interaction  $\lambda$ , hence  $\lambda_{eff} = \lambda + \lambda_{mag}$  is negative, no matter what is the sign of  $\lambda$ .

One can go a step further and add to the interaction vertex in Fig. 8 (the shaded rectangle) the renormalizations coming from fermions with energies higher than  $\Omega_{mag}$ . These terms renormalize the convolution of two spin-fluctuation propagators  $\chi^2(\Omega)$  into  $\chi_{I-nem} = \chi^2/[1 + 2g\chi^2(\Omega)]$  and hence add the same denominator to  $\lambda_{mag}$ . As the consequence,  $\lambda_{mag}$  diverges already at the Ising-nematic instability, before  $m_s$  vanishes.

By continuity, we assume that  $\lambda_{mag}$  remains negative also for finite hole and electron pockets. Using further RPA form for the Raman intensity, Eq. (2), we obtain that  $\chi_R(\Omega)$  has the singularity at  $\Omega = \Omega_{res,2} < 2\Delta$ , at which  $1 + \lambda_{eff}\Pi_{B_{1g}}(\Omega) = 0$ .

Finally, we briefly comment on the difference between our analysis and earlier phenomenological consideration of the bi-linear coupling between  $B_{1g}$  orbital order parameter  $\Delta_{oo} = \sum_k < c_k^\dagger c_k \cos 2\theta_k >$  and  $\Delta_1^2 - \Delta_2^2$ , where  $\Delta_1$  and  $\Delta_2$  are spin-fluctuation fields with momenta near  $\mathbf{Q}_1$  and  $\mathbf{Q}_2$  [13 and 23]. In the microscopic calculation [28] such term appears in the Landau free energy once we introduce  $\Delta_{1,2}$  and  $\Delta_{oo}$  as order parameter fields, bi-linear in fermions, and perform Hubbard-Stratonovich transformation from fermions to bosonic collective variables. The prefactor  $A$  for  $\Delta_{oo}(\Delta_1^2 - \Delta_2^2)$  term in the Landau functional is given by the same triangular diagram as AL vertex, and has a finite value (i.e.,  $A \sim \Gamma_{tr}$ ). At the first glance, we can identify  $\Delta_1^2 - \Delta_2^2$  with the propagator of an Ising-nematic field and obtain the correction to the prefactor for  $\Delta_{oo}^2$  in the form  $-A^2\chi_{I-nem}$ . Because the bare prefactor is  $\Pi_{B_{1g}}^{-1} + \lambda$ ,  $\lambda_{eff} = \lambda - A^2\chi_{I-nem}$ . At the second glance, however, we note that the Landau functional in terms of  $\Delta_1$  and  $\Delta_2$  is not the same as Landau functional expressed in terms of the Ising-nematic field. To obtain the latter one has to do a second Hubbard-Stratonovich transformation to the composite Ising-nematic bosonic field  $\Delta_{I-nem}$  and integrate over the primary fields  $\Delta_1$  and  $\Delta_2$ . Only then one can extract the bi-linear coupling between orbital and Ising-nematic order parameters. Another way to see that  $A^2\chi_{I-nem}$  with  $A \sim \Gamma_{tr}$  is not the correction to  $\lambda$  is to notice that this expression is the full result for the Raman bubble

rather than for the effective interaction between fermions from the two particle-hole bubbles.

### C. Comparative analysis of the two scenarios

Combining the results of the last two sections, we see that the resonance in  $\chi_R(\Omega)$  holds independent of whether  $\Omega_{mag}$  is larger or smaller than  $\Delta$ , but the physics is different in the two cases. When  $\Omega_{mag}$  is smaller than  $\Delta$ , the resonance has purely magnetic origin and comes from the pole in  $\chi_{I-nem}$  at  $\Omega = \Omega_{res,1} \leq 2\Omega_{mag}$ . For this resonance, the role of fermions is to provide some regular coupling,  $\Gamma_{tr}$ , between incoming and outgoing light and a pair of spin fluctuations with momenta near  $\mathbf{Q}_1$  or  $\mathbf{Q}_2$ . When  $\Omega_{mag} > \Delta$ , the resonance comes from fermions and is due to singular behavior of particle-hole polarization bubble  $\Pi_{B_{1g}}(\Omega)$  at  $\Omega = 2\Delta$ . The resonance occurs at a frequency  $\Omega = \Omega_{res,2} \leq 2\Delta$ . Spin fluctuations are again crucial, but now their role is to provide strong attractive interaction between fermions which make particle-hole bubbles.

We treated the two singularities in  $\chi_R(\Omega)$  independent of each other chiefly to demonstrate that they come from two different pieces of physics. Such a treatment, however, is justified only if  $\Omega_{res,1}$  and  $\Omega_{res,2}$  are well separated. In our case,  $\Omega_{res,1} \leq 2\Omega_{mag} < 4\Delta$ , while  $\Omega_{res,2} \leq 2\Delta$ . How well  $\Omega_{res,1}$  and  $\Omega_{res,2}$  are separated then depends on the strength of various interactions and on the value of magnetic correlation length. Like we said, in FeSCs that we analyze,  $\Omega_{mag}$  and  $\Delta$  are not far from each other. In this case the resonance likely has a dual origin. In NaFe<sub>1-x</sub>Co<sub>x</sub>As with  $x = 0.045$ , the  $B_{1g}$  peak is seen at 7.1 meV, which is below both  $2\Delta$  and  $2\Omega_{mag}$ . This is consistent with the dual origin of the resonance.

### D. Numerical evaluation of the AL diagrams

In this section we present the results of numerical evaluation of AL contributions to  $B_{1g}$  Raman intensity below  $T_c$  first in the two-loop approximation and then including the interaction between spin fluctuations. We first compute the AL vertex assuming that  $\Gamma_{tr}$  can be approximated by a constant and then present the results of the explicit calculation of the two-loop AL diagram for  $\chi_R(\Omega)$ .

The first step for numerical evaluation of  $\chi_R$  is to calculate the time-ordered polarization function  $\Pi_{s,ij}$  with fermions lines from bands  $i$  and  $j$ . The bare spin response can be obtained as a time-ordered average of spin operators over a non-interacting ground state. In the FeSCs, the response is largest near the nesting momenta  $\mathbf{Q}_1 = (\pi, 0)$  or  $\mathbf{Q}_2 = (0, \pi)$ , which connect one hole pocket and one electron pocket. Since we are solely interested in evaluating the function near those momenta we will only consider band combinations of one hole and one electron pocket and drop the band indices from here

on. We evaluate the function at momentum  $\mathbf{Q} + \mathbf{q}$ , where  $\mathbf{Q}$  is either  $\mathbf{Q}_1$  or  $\mathbf{Q}_2$ , whichever is appropriate.

For concreteness, we assume parabolic dispersions for the hole and electron pockets of the form  $\xi_{\mathbf{k}} = \mu - \frac{k^2}{2m_h}$  and  $\xi_{\mathbf{k}+\mathbf{Q}} = -\mu + \frac{k_x^2}{2m_x} + \frac{k_y^2}{2m_y}$ , respectively. We evaluate all quantities in units of the gap  $\Delta$  and for numerical parameters we choose  $\mu = 2\Delta$ ,  $m_h \approx 0.056\Delta^{-1}$  ( $k_F = 0.15\pi/a$ ),  $m_x = m_h/1.27$ , and  $m_y = m_x/0.3787$ . These values approximately fit the bands and Fermi surfaces reported in ARPES measurements [41] of NaFe<sub>1-x</sub>Co<sub>x</sub>As for  $x = 0.05$  (of the two hole bands, we fitted the one with the largest Fermi surface). For numerical convergence we included a finite broadening  $\eta = \Delta/100$ .

The general behavior of  $\Pi_s$  can be seen in Figs. 11 and 12. The first one shows a frequency sweep of the real part at  $\mathbf{q} = 0$  and the divergence at  $2\Delta$  is clearly seen. The imaginary part (not shown) vanishes as  $\eta \rightarrow 0$ . This behavior holds unless  $\mathbf{q}$  is so large that the normal-state FSs no longer intersect due to the shift. The second plot shows an example of the  $\mathbf{q}$  dependence at  $\nu = 0$ . Although the function is anisotropic due to the eccentricity of the electron Fermi surface, the qualitative behavior is the same regardless of the polar angle. It is particularly important to emphasize that the function decreases monotonically with increasing  $|\mathbf{q}|$ .

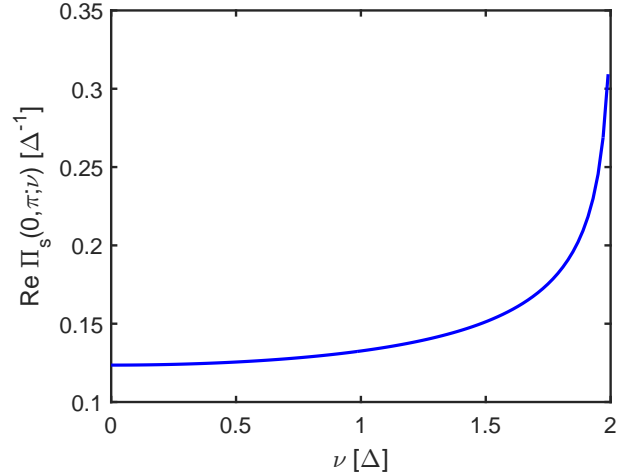


Figure 11. Frequency dependence of the bare spin polarization operator with momentum  $(0, \pi)$ .

In the numerical analysis, it is easier to deal with the effective interaction in the spin channel  $U_{eff}$  rather than the spin susceptibility. In the RPA

$$U_{eff}(\mathbf{Q} + \mathbf{q}, \nu) = \frac{u}{2} \frac{1}{1 - u\Pi_s(\mathbf{Q} + \mathbf{q}, \nu)}, \quad (31)$$

where  $u > 0$  is the bare fermion-fermion interaction ( $= U$  in the Hubbard model). The effective interaction is related to the spin susceptibility by  $U_{eff}(\mathbf{Q} + \mathbf{q}, \nu) = [u + u^2\chi^s(\mathbf{Q} + \mathbf{q}, \nu)]/2$ , so the two functions have the same pole structure and differ only by a constant shift

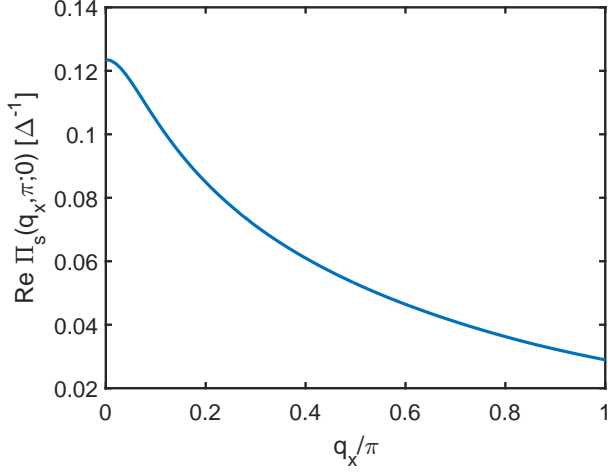


Figure 12. Momentum dependence of the static spin polarization operator with momentum near  $(0, \pi)$ . We have chosen momenta to be  $\mathbf{q} + (0, \pi)$ . The dependence on  $q_x$  is shown. The dependence on the polar angle of  $\mathbf{q}$  is nonzero but rather weak.

$u$ , which near a magnetic instability is small compared to the second term. The  $u^2/2$  factor in front of  $\chi^s$  is the same factor as in (9).

Figure 13 shows  $U_{eff}$  at  $\mathbf{q} = 0$  as a function of  $\nu$ . The real part diverges at the resonance frequency while the imaginary part has a sharp peak which in the limit of  $\eta \rightarrow 0$  becomes a  $\delta$  function. For the numerical calculations we have set  $u \approx 7.9\Delta$ , which determines  $\Omega_{mag} \approx 0.6\Delta$ .

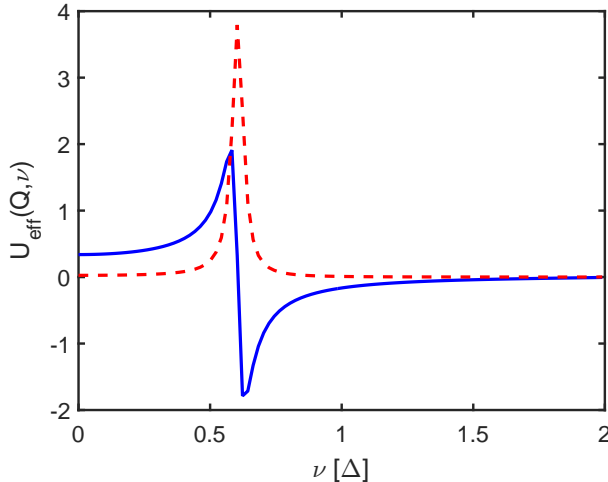


Figure 13. Frequency dependence of the real (solid blue line) and imaginary (dashed red line) parts of the effective interaction (in arbitrary units).

Now we are ready to evaluate the Raman response (20). By using the spectral representation, the imaginary part of the response function can be equivalently calculated

as

$$\text{Im } \chi_R(\Omega) \propto \int \frac{d^2 \mathbf{q}}{(2\pi)^2} \int_0^\Omega \frac{d\nu}{\pi} \text{Im } U_{eff}(\mathbf{Q} + \mathbf{q}, \nu) \times \text{Im } U_{eff}(\mathbf{Q} + \mathbf{q}, \nu - \Omega). \quad (32)$$

The advantage of this form is that it only requires knowledge of the function in a finite range of  $\nu$ . The real part can then be calculated by using the *KK* transformation.

Because  $\text{Im } U_{eff}(\mathbf{Q} + \mathbf{q}, \nu)$  is peaked at the resonant frequencies corresponding to each momentum  $\mathbf{q}$ ,  $\text{Im } \chi_R(\Omega)$  can be seen as a convolution of many of these peaks. The result of the computation is shown in Fig. 14. We see that the imaginary part starts small and undergoes a jump at  $\nu \approx 1.2\Delta = 2\Omega_{mag}$ , (compare to Fig. 13). Its value at higher frequencies comes from contributions from  $\mathbf{q} \neq 0$ , corresponding to excitons of higher energies. This jump is not sharp in the numerical calculations because of the finite value of  $\eta$  in (17).

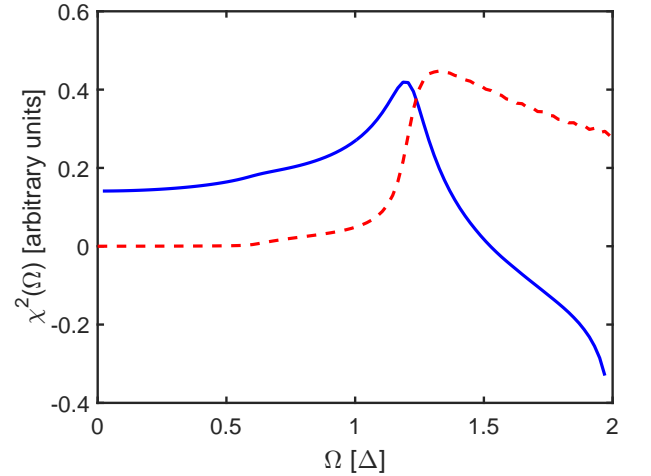


Figure 14. Contribution to Raman response from spin fluctuations. The solid blue and dashed red lines indicate the real and imaginary parts, respectively.

We next consider the effect of the finite state interaction, i.e., include higher order diagrams (Fig. 3) with the interactions between the two spin fluctuations. In the approximation where the interactions between pairs of spin fluctuations with momenta near  $\mathbf{Q}_1$  and/or  $\mathbf{Q}_2$  can be treated as constants, we use Eq. (24). We plot the RPA form of  $\chi_R(\Omega)$  in Fig. 15. We clearly see that  $\text{Im } \chi_R^0(\Omega)$  has a sharp peak at a frequency  $\Omega < 2\Omega_{mag}$ , which is below  $2\Delta$ .

## V. TRIANGULAR FERMION LOOP AND SYMMETRY CHANNELS

So far, we have neglected the details of the Raman vertex  $\Gamma_{tr}$  and thus the analysis above applies equally to all symmetry channels. We remind that in the experiments the resonance has been observed in the  $B_{1g}$  channel in

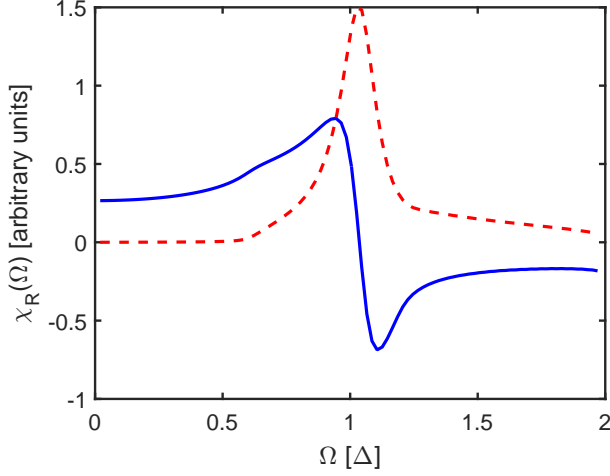


Figure 15. The full Raman susceptibility, which includes interactions between spin fluctuations. The solid blue line indicates the real part and the dashed red line the imaginary part of the function. We set  $g = 1.66$ .

the 1-Fe unit cell, but no resonance has been observed in the  $B_{2g}$  or  $A_{1g}$  channels. To address the origin of the difference between Raman scattering in different symmetry channels, we consider the symmetry properties of the function  $\Gamma_{tr}(\mathbf{q}, \nu)$ , defined in (7).

First we focus on the bare Raman vertices. The original Raman vertex  $\gamma_i(\mathbf{k})$  between incoming and outgoing light and incoming and outgoing fermions from band  $i$  belongs to the irreducible representations of the point group  $D_{4h}$ . As a reminder, we point out that in the  $B_{1g}$  and  $B_{2g}$  representations, the function  $\gamma_i(\mathbf{k})$  transforms under the point group operations as  $k_x^2 - k_y^2$  and  $k_x k_y$ , respectively, while in the  $A_{1g}$  representation  $\gamma_i(\mathbf{k})$  is invariant under the point group operations. We will consider each of these channels separately.

In the  $B_{1g}$  channel, the Raman response is directly coupled to nematic orbital fluctuations and in the orbital basis we can define it as

$$\chi_R(\Omega) = -i \int dt e^{i\Omega t} \langle T \rho_n(t) \rho_n(0) \rangle, \quad (33)$$

where the nematic "density" operator is defined as  $\rho_n = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}, \sigma} [a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} - b_{\mathbf{k}\sigma}^\dagger b_{\mathbf{k}\sigma}]$ . In this notation, the fermion operators  $a$  and  $b$  correspond to the  $d_{xz}$  and  $d_{yz}$  orbitals, respectively.

The transformation from the orbital to the band basis in the case of the hole pockets can be approximated by [50]

$$\begin{aligned} \alpha_{\mathbf{k}\sigma} &= \cos \theta_{\mathbf{k}} a_{\mathbf{k}\sigma} - \sin \theta_{\mathbf{k}} b_{\mathbf{k}\sigma}, \\ \beta_{\mathbf{k}\sigma} &= \sin \theta_{\mathbf{k}} a_{\mathbf{k}\sigma} + \cos \theta_{\mathbf{k}} b_{\mathbf{k}\sigma}, \end{aligned} \quad (34)$$

where  $\alpha$  and  $\beta$  denote the hole bands and  $\theta$  is the angle along the Fermi surface.

The contribution to the nematic density operator from

fermions from the hole bands then becomes

$$\begin{aligned} \rho_n &= \sum_{\mathbf{k}} (\alpha_{\mathbf{k}\sigma}^\dagger \alpha_{\mathbf{k}\sigma} - \beta_{\mathbf{k}\sigma}^\dagger \beta_{\mathbf{k}\sigma}) \cos 2\theta_{\mathbf{k}} \\ &+ \sum_{\mathbf{k}} (\alpha_{\mathbf{k}\sigma}^\dagger \beta_{\mathbf{k}\sigma} + \beta_{\mathbf{k}\sigma}^\dagger \alpha_{\mathbf{k}\sigma}) \sin 2\theta_{\mathbf{k}}. \end{aligned} \quad (35)$$

The second term can be neglected at low energies because it couples fermions from different Fermi surfaces, which do not cross. Substituting into (33) we find that the vertex function for holes  $\gamma_i(\mathbf{k})$  in the  $B_{1g}$  channel is  $\cos(2\theta_{\mathbf{k}})$  and has opposite signs for the two hole bands. For electron pockets, the situation is more simple since each electron pocket only has contributions from the  $d_{zx}$  or  $d_{yz}$  orbital, but not from both. The transformation from the orbital to the band basis is trivial and we find that  $\gamma_i(\mathbf{k}) = \pm 1$ , where the plus sign is for the electrons from the pocket near  $\mathbf{Q}_1 = (\pi, 0)$  and the minus sign for electrons near  $\mathbf{Q}_2 = (0, \pi)$ .

For the  $B_{2g}$  channel one can define a different density operator  $\rho_{B2g} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}, \sigma} [a_{\mathbf{k}\sigma}^\dagger b_{\mathbf{k}\sigma} + b_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma}]$ . Then using again the transformation (34) one finds that for hole bands the appropriate  $B_{2g}$  symmetry factor is  $\gamma_i(\mathbf{k}) = \pm \sin(2\theta_{\mathbf{k}})$ . For electron bands the symmetry factor is instead  $\gamma_i(\mathbf{k}) = 0$  since the electron bands do not cross.

Finally, the Raman vertex in the  $A_{1g}$  channel is a constant along hole or electron pockets. In general, there are two possibilities: The symmetry factor can have the same sign for the coupling of light to electrons and holes, or change sign when switching between electrons and holes. In order to consider both possibilities we define two separate functions  $\gamma_{A1g}^{++}$  and  $\gamma_{A1g}^{+-}$ , referring to the cases with equal and opposite signs, respectively.

An additional effect of the orbital to band transformation (34) is that the factors of sine and cosine contribute extra angular dependence to the momentum integration in  $\Lambda$ . This is summarized graphically in Fig. 16, where we list the different band combinations with the appropriate signs and angular dependences. The symmetry factor  $\gamma(\mathbf{k})$  for each diagram in different symmetry channels is given in Table I. The first four diagrams [Figs. 16(a)-(d)] are for the interaction between light and spin fluctuations with momentum near  $\mathbf{Q}_1$  and the other four are for momentum near  $\mathbf{Q}_2$ . The total contribution to  $\Lambda$  in each case is given by the sum of the four diagrams. The angular dependencies listed in the figure are for a model with only intra-orbital Hubbard interaction.

The result of the momentum integration is different depending on the symmetry channel. For simplicity in the evaluation of the integral, we consider identical circular Fermi surfaces in all bands and evaluate the bare triangular vertices  $\Gamma_{tr}^0(\mathbf{q}, \nu)$  in various geometries at  $\mathbf{q} = 0$ . This particular value of  $\mathbf{q}$  is important because the enhancement in  $\chi^2(\Omega)$  at  $\Omega = \Omega_{mag}$  comes primarily from momenta near  $\mathbf{q} = 0$ . Considering only the angular part



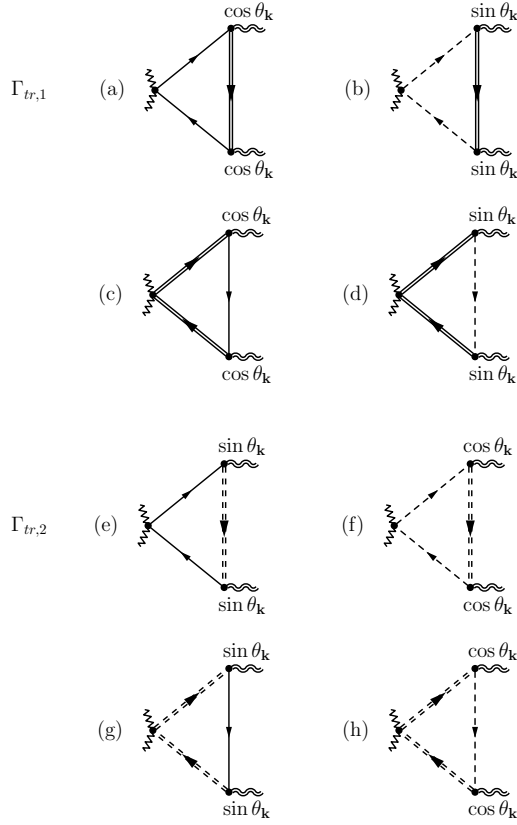


Figure 16. Contributions to the triangular fermion loop from different bands. The single solid and dashed lines represent the hole bands  $\alpha$  and  $\beta$ , respectively, while the double solid and dashed lines represent electron bands centered at momentum  $\mathbf{Q}_1$  and  $\mathbf{Q}_2$ , respectively. The factors of  $\cos \theta$  and  $\sin \theta$  arise from the transformation from the orbital to the band basis (see text).

of the integration, we find that

$$\Gamma_{tr}^{0,B_{1g}} \propto \int d\theta_{\mathbf{k}} (\cos 2\theta_{\mathbf{k}} \cos 2\theta_{\mathbf{k}} - 1) \neq 0 \quad (36)$$

$$\Gamma_{tr}^{0,B_{2g}} \propto \int d\theta_{\mathbf{k}} \sin 2\theta_{\mathbf{k}} \cos 2\theta_{\mathbf{k}} = 0 \quad (37)$$

$$\Gamma_{tr}^{0,++} \propto \int d\theta_{\mathbf{k}} (1 - 1) = 0 \quad (38)$$

$$\Gamma_{tr}^{0,+ -} \propto \int d\theta_{\mathbf{k}} (1 + 1) \neq 0 \quad (39)$$

where, we remind,  $\Gamma_{tr}^{0,++}$  and  $\Gamma_{tr}^{0,+ -}$  are two different triangular vertices in  $A_{1g}$  channel. We see that  $\Gamma_{tr}^{0,B_{2g}}$  and  $\Gamma_{tr}^{0,++}$  vanish, i.e., there is no enhancement of the Raman intensity in the  $B_{2g}$  or  $A_{1g}^{++}$  channels, in agreement with the data.

One fine point in the calculation is that the contribution from diagrams with electron bands at the vertex has an additional minus sign compared to diagrams with hole bands at the vertex [for example, compare diagrams Figs. 16(c) and (d) with Figs. 16(a) and (b)]. This is

Diagram	$\gamma_{B_{1g}}$	$\gamma_{B_{2g}}$	$\gamma_{A_{1g}}^{++}$	$\gamma_{A_{1g}}^{+-}$
Fig. 16(a)	$\cos 2\theta$	$\sin 2\theta$	1	1
Fig. 16(b)	$-\cos 2\theta$	$-\sin 2\theta$	1	1
Fig. 16(c)	1	0	1	-1
Fig. 16(d)	1	0	1	-1
Fig. 16(e)	$\cos 2\theta$	$\sin 2\theta$	1	1
Fig. 16(f)	$-\cos 2\theta$	$-\sin 2\theta$	1	1
Fig. 16(g)	-1	0	1	-1
Fig. 16(h)	-1	0	1	-1

Table I. Symmetry factors for Raman vertices in different symmetry channels. We include two different representations of  $A_{1g}$  symmetry: One where the sign changes between hole and electron pockets, and one in which it does not (see text).

not a symmetry factor but instead comes from the opposite signs between the hole and electron band dispersions and can be obtained after performing the frequency integration in the fermion loop. In the sign-preserving  $A_{1g}$  channel this leads to a complete cancellation while in the  $B_{1g}$  channel it is only partial since, e.g., Figs. 16(a) and (b) contribute a factor of  $\cos 2\theta$  while Figs. 16(c) and (d) contribute a factor of  $-\frac{1}{2} \cos 2\theta$ .

To summarize, we have shown so far that the bare vertices for the  $B_{2g}$  and sign-preserving  $A_{1g}$  channels vanish and thus cannot lead to resonances. The bare vertices in  $B_{1g}$  and sign-changing  $A_{1g}$  channels are nonzero. Note that the Raman response in the sign-changing  $A_{1g}$  channel is not screened out by long-range component of Coulomb interaction [3 and 4]. The response in the sign-preserving  $A_{1g}$  channel is screened [15–17], but in our case it vanishes anyway.

The next step is to include interactions between the spin fluctuations and calculate the renormalized vertices  $\Gamma_{tr}^{B_{1g}}$  and  $\Gamma_{tr}^{+-}$ .

For this we first introduce vertices  $\Gamma_{tr,1}$  and  $\Gamma_{tr,2}$ , which couple light to spin fluctuations with momentum near  $\mathbf{Q}_1$  and  $\mathbf{Q}_2$ , respectively. The vertices in  $B_{1g}$  and  $A_{1g}$  channels are related to  $\Gamma_{tr,1}$  and  $\Gamma_{tr,2}$  as

$$\Gamma_{tr,1} = \Gamma_{tr}^{B_{1g}} + \Gamma_{tr}^{+-}, \quad \Gamma_{tr,2} = \Gamma_{tr}^{+-} - \Gamma_{tr}^{B_{1g}} \quad (40)$$

We then follow Ref. [28] and model the interaction between spin fluctuations as given by the effective action

$$S_{eff} = r_0(\Delta_1^2 + \Delta_2^2) + \frac{\kappa}{2}(\Delta_1^2 + \Delta_2^2)^2 + \frac{g}{2}(\Delta_1^2 - \Delta_2^2)^2, \quad (41)$$

where  $\Delta_1$  and  $\Delta_2$  are three-component spin fluctuation fields, respectively, in which each component corresponds to a direction in real space. The 1 and 2 subscripts distinguish between fluctuations near  $\mathbf{Q}_1$  and  $\mathbf{Q}_2$ . This effective action can be obtained by introducing Hubbard-Stratonovich fields and then integrating out fermions. The result is [28] that  $\kappa > 0$ , but  $g$  is negative, at least at small dopings.

The bare vertices  $\Gamma_{tr,1}^0$  and  $\Gamma_{tr,2}^0$  are given by the sum of diagrams Figs. 16(a)-(d) and Figs. 16(e)-(h), respectively. In the ladder approximation, the coupled equations for the renormalized vertices  $\Gamma_{tr,1}$  and  $\Gamma_{tr,2}$  are given by

$$\Gamma_{tr,1} = \Gamma_{tr,1}^0 - (\kappa + g)\Gamma_{tr,1}\chi^2 - (\kappa - g)\Gamma_{tr,2}\chi^2, \quad (42)$$

$$\Gamma_{tr,2} = \Gamma_{tr,2}^0 - (\kappa + g)\Gamma_{tr,2}\chi^2 - (\kappa - g)\Gamma_{tr,1}\chi^2, \quad (43)$$

where we remind that  $\chi^2$  is the shorthand notation for  $\int \frac{d^2\mathbf{q}d\nu}{(2\pi)^3} \chi^s(\mathbf{Q} + \mathbf{q}, \nu) \chi^s(\mathbf{Q} + \mathbf{q}, \nu + \Omega)$ . In this expression we do not distinguish between  $\mathbf{Q}_1$  and  $\mathbf{Q}_2$  because by symmetry the result of the integration is the same.

These coupled equations can be solved in terms of  $\Gamma_{tr}^{B_{1g}}$  and  $\Gamma_{tr}^{+-}$ :

$$\Gamma_{tr}^{+-} = \frac{\Gamma_{tr}^{0,+ -}}{1 + 2\kappa\chi^2}, \quad (44)$$

$$\Gamma_{tr}^{B_{1g}} = \frac{\Gamma_{tr}^{0,B_{1g}}}{1 + 2g\chi^2}. \quad (45)$$

The renormalized Raman susceptibility is then given by

$$\chi_R^{B_{1g}} = \frac{(\Gamma_{tr}^{0,B_{1g}})^2 \chi^2}{1 + 2g\chi^2}, \quad (46)$$

$$\chi_R^{A_{1g},+-} = \frac{(\Gamma_{tr}^{0,+ -})^2 \chi^2}{1 + 2\kappa\chi^2}. \quad (47)$$

Since  $\kappa > 0$  and  $g < 0$ , only the  $B_{1g}$  channel has a resonance, which is consistent with the data. There is no resonance-type enhancement from the coupling to spin fluctuations, regardless of whether the sign-preserving or sign-changing representation is involved. We note in passing that there is a different enhancement of the Raman intensity in the  $A_{1g}$  channel in  $s^{+-}$  superconductors due to a direct process in which a light generates a ladder series of particle-hole pairs [3].

We also note that the downward renormalization of  $\Gamma_{tr}^{+-}$  by  $1/(1 + 2\kappa\chi^2)$  also strongly reduces the temperature dependence of the  $A_{1g}$  Raman intensity in the normal state. Indeed, explicit calculation shows [28] that  $\kappa \sim 1/T^2$ . Following the considerations of Sec. III, we find that in the normal state

$$\chi_R^{A_{1g}}(T) \sim \frac{(m_s)^{-2}}{T + T_1}, \quad (48)$$

where  $T_1 \gg T_0$  because  $\kappa \gg |g|$ . This implies that at  $T \geq T_0$ ,  $\chi_R^{A_{1g}}(T)$  is nearly flat.

## VI. CONCLUSIONS

In this work we argued that the coupling of the Raman vertex to pairs of magnetic fluctuations via the AL process can explain the  $1/(T - T_0)$  behavior of  $B_{1g}$  Raman intensity in the normal state of  $\text{NaFe}_{1-x}\text{Co}_x\text{As}$ ,  $\text{EuFe}_2\text{As}_2$ ,

$\text{SrFe}_2\text{As}_2$ , and  $\text{Ba}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$  and the development of the resonance below  $T_c$ , observed in  $\text{NaFe}_{1-x}\text{Co}_x\text{As}$  and  $\text{Ba}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$ .

We considered the AL process in which light couples to a particle-hole pair, which then gets converted into a pair of spin fluctuations with momenta near  $\mathbf{Q}_1 = (\pi, 0)$  and  $\mathbf{Q}_2 = (0, \pi)$ . We analyzed magnetically mediated Raman intensity both analytically and numerically, first at the leading two-loop order and then included interactions between pairs of magnetic fluctuations. We demonstrated explicitly that the full Raman intensity in the  $B_{1g}$  channel can be viewed as the result of the coupling of light to Ising-nematic susceptibility via Aslamazov-Larkin process. We argued that the  $1/(T - T_0)$  temperature dependence in the normal state is the combination of the temperature dependencies of the Aslamazov-Larkin vertex and of Ising-nematic susceptibility. We further argued that the resonance in the  $B_{1g}$  channel below  $T_c$  emerges because of two effects. One is the development of a pole in the fully renormalized Ising-nematic susceptibility. The pole occurs at a frequency  $\Omega_{res,1} < 2\Omega_{mag}$ , where  $\Omega_{mag}$  is the minimal frequency of a dispersing spin resonance at momenta near  $Q_{1,2}$  in an  $s^{+-}$  superconductor. Another effect is that spin fluctuations generate attractive interactions between low-energy fermions, which constitute particle-hole bubbles with zero momentum transfer. An attractive interaction between such fermions combined with the fact that in an  $s$ -wave superconductor a particle-hole bubble at zero momentum transfer is singular at  $2\Delta$  gives rise to an excitonic resonance at  $\Omega_{res,2} < 2\Delta$ . In FeSCs  $\Omega_{res,1}$  and  $\Omega_{res,2}$  are not far from each other, and the observed strong peak in  $B_{1g}$  Raman intensity below  $T_c$  is likely the mixture of both effects.

## VII. ACKNOWLEDGEMENTS

We acknowledge with thanks conversations with G. Blumberg, Y. Gallais, R. Hackl, R. Fernandes, M. Khodas, H. Kontani, I. Paul, A. Sacuto, J. Schmalian, V. Thorsmølle, and R. Xing. The work is supported by the DOE Grant No. de-sc0014402.

## Appendix A: Orbital fluctuations

In this Appendix we consider the coupling of the Raman response to orbital fluctuations in detail and show that the interaction in the  $d$ -wave channel is repulsive and cannot lead to the observed resonance.

As explained in the Introduction, the electronic structure of  $\text{NaFe}_{1-x}\text{Co}_x\text{As}$  and  $\text{Ba}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$  consists of four bands that cross the Fermi energy. We will refer to the two hole bands centered at  $(0, 0)$  as  $\alpha$  and  $\beta$  and to the electron bands centered at  $(\pi, 0)$  and  $(0, \pi)$  as  $\eta$  and  $\delta$ , respectively. We organize our analysis in the language of vertex renormalization. We start with a set of bare Raman vertices  $\gamma_i(\mathbf{k})$  with both external fermion

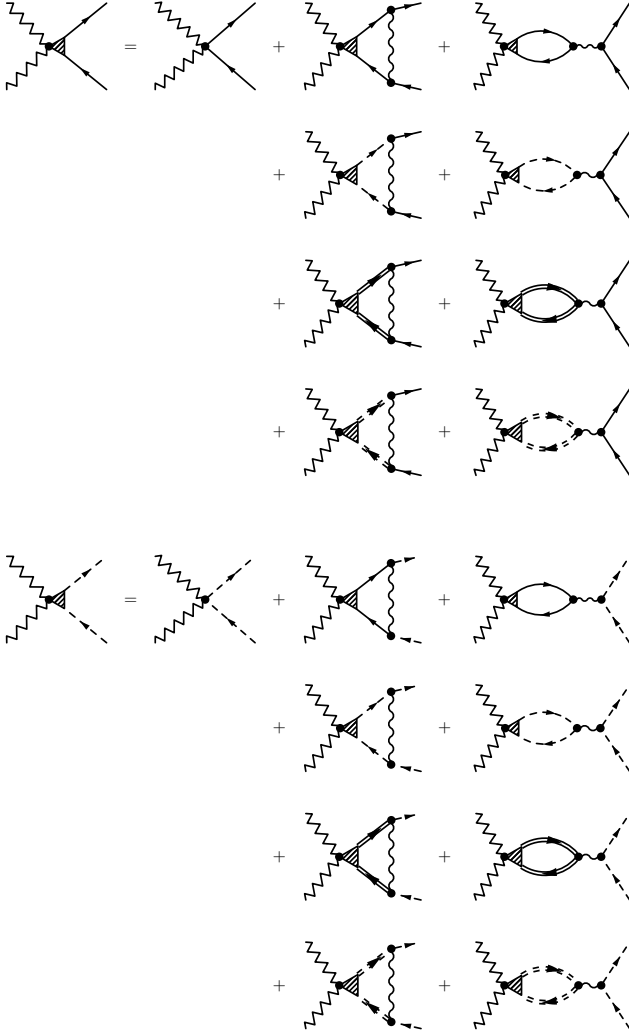


Figure 17. Vertex renormalization for hole bands  $\alpha$  and  $\beta$ . The single solid and dashed lines represent excitations from hole bands  $\alpha$  and  $\beta$ , respectively, and the double solid and dashed lines from electron bands  $\eta$  and  $\delta$ , respectively.

lines belonging to the  $i$ -th band, then dress each vertex with interactions to obtain the full vertex  $\Gamma_i(\mathbf{k})$ . Here the index runs over the set of bands  $\{\alpha, \beta, \eta, \delta\}$ . Since we are interested in computing the Raman response in the limit of vanishingly small external momentum, we do not consider vertices with external fermion lines from two different bands since in the absence of band crossings there will be no low-energy contribution from these vertices.

For simplicity we study a model consisting of only  $d_{xz}$  and  $d_{yz}$  orbitals. Since the remaining  $d_{xy}$  orbital has  $A_{1g}$  symmetry it cannot directly contribute to the Raman response in the  $B_{1g}$  channel. The transformation between the band and orbital basis for the hole pockets was given in (34). For the electron pockets we will simply set  $\eta_{\mathbf{k}+\mathbf{Q}_1} = a_{\mathbf{k}+\mathbf{Q}_1}$  and  $\delta_{\mathbf{k}+\mathbf{Q}_2} = b_{\mathbf{k}+\mathbf{Q}_2}$ . Following the reasoning in Sec. V, the bare  $B_{1g}$  Raman vertex is

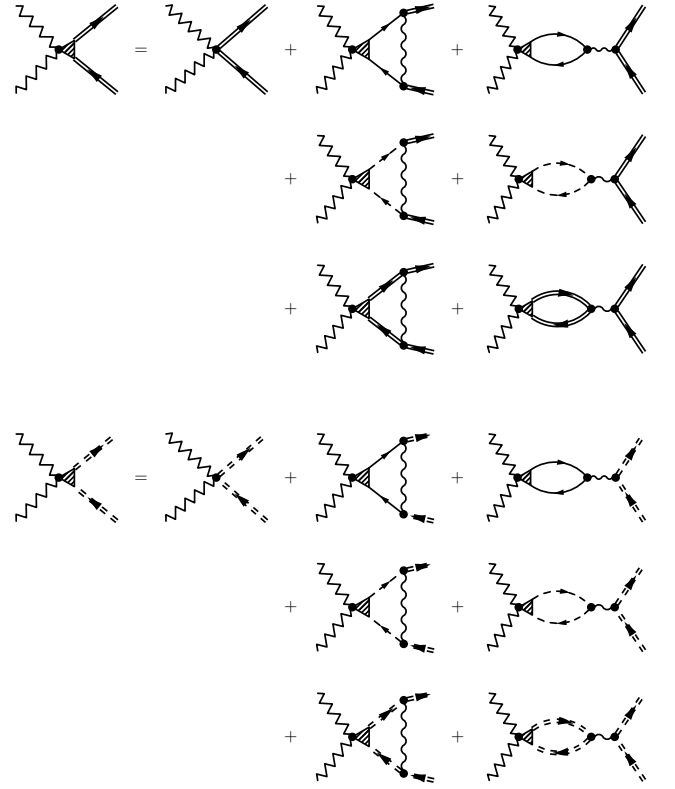


Figure 18. Vertex renormalization for electron bands  $\eta$  and  $\delta$ . The notations are the same as in Fig. 17.

given by  $\gamma_i(\mathbf{k}) = \{\cos 2\theta_{\mathbf{k}}, -\cos 2\theta_{\mathbf{k}}, 1, -1\}_i$ . The alternating signs reflect the difference between  $d_{xz}$  and  $d_{yz}$  contributions.

For our perturbative analysis, we assume that the short-range intra-orbital repulsion is the dominant interaction. Thus, the interaction Hamiltonian in the orbital basis is given by

$$\mathcal{H}_I = U \sum_{\mathbf{q}} [\rho_{xz}(\mathbf{q})\rho_{xz}(-\mathbf{q}) + \rho_{yz}(\mathbf{q})\rho_{yz}(-\mathbf{q})], \quad (\text{A1})$$

where  $\rho_{xz}(\mathbf{q}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}, \sigma} a_{\mathbf{k}+\mathbf{q}\sigma}^\dagger a_{\mathbf{k}\sigma}$  and  $\rho_{yz}(\mathbf{q}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}, \sigma} b_{\mathbf{k}+\mathbf{q}\sigma}^\dagger b_{\mathbf{k}\sigma}$  are the density operators of the  $d_{xz}$  and  $d_{yz}$  orbitals, respectively.

Our diagrammatic analysis is summarized in Figs. 17 and 18. For notational convenience we define auxiliary functions  $\tilde{\Gamma}_i(\mathbf{k})$  and  $\tilde{\gamma}_i(\mathbf{k})$  such that  $\Gamma_i = \{\tilde{\Gamma}_\alpha \cos 2\theta_{\mathbf{k}}, \tilde{\Gamma}_\beta \cos 2\theta_{\mathbf{k}}, \tilde{\Gamma}_\eta, \tilde{\Gamma}_\delta\}_i$  and a similar expression for  $\gamma_i$ . The set of coupled equations for the Raman vertices can be written in matrix form as

$$\tilde{\mathbf{\Gamma}} = \tilde{\boldsymbol{\gamma}} - \mathbf{V}\mathbf{\Pi}\tilde{\boldsymbol{\gamma}}, \quad (\text{A2})$$

where  $\mathbf{V}$  and  $\mathbf{\Pi}$  are interaction and polarization matrices,



respectively, given by

$$\mathbf{V} = \frac{U}{2} \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 2 & 0 \\ -1 & 1 & 0 & 2 \end{pmatrix} \quad (\text{A3})$$

$$\mathbf{\Pi} = \begin{pmatrix} \Pi_\alpha(0, \Omega) & 0 & 0 & 0 \\ 0 & \Pi_\beta(0, \Omega) & 0 & 0 \\ 0 & 0 & \Pi_\eta(0, \Omega) & 0 \\ 0 & 0 & 0 & \Pi_\delta(0, \Omega) \end{pmatrix} \quad (\text{A4})$$

In this notation, the polarization functions are defined as

$$\begin{aligned} \Pi_i(0, \Omega) = & i \int \frac{d^2 k d\nu}{(2\pi)^3} s_i(\mathbf{k}) [G_i(\mathbf{k}, \nu + \Omega) G_i(\mathbf{k}, \nu) \\ & - F_i(\mathbf{k}, \nu + \Omega) F_i(\mathbf{k}, \nu)], \end{aligned} \quad (\text{A5})$$

where  $G_i$  and  $F_i$  are normal and anomalous Green's functions for band  $i$ , respectively, and  $s_i(\mathbf{k}) = \{\cos^2(2\theta_{\mathbf{k}}), \cos^2(2\theta_{\mathbf{k}}), 1, 1\}_i$ . We note that by symmetry  $\Pi_\eta(0, \Omega) = \Pi_\delta(0, \Omega)$ . In this definition, the real part of each function is positive.

The solution to (A2) is easily obtained as  $\tilde{\Gamma} = (\mathbb{I} + \mathbf{V}\mathbf{\Pi})^{-1}\tilde{\gamma}$ , where  $\mathbb{I}$  is the identity matrix. After evaluation of the matrix multiplication we find that the full vertex is given by

$$\Gamma_i = \left[ 1 + \frac{U}{2} \sum_j \Pi_j(0, \Omega) \right]^{-1} \gamma_i \quad (\text{A6})$$

and the full response function takes the simple form

$$\chi_R(\Omega) = \frac{2 \sum_i \Pi_i(0, \Omega)}{1 + \lambda \sum_j \Pi_j(0, \Omega)}, \quad (\text{A7})$$

where  $\lambda = U/2 > 0$ . This is the same formula as Eq. (2) [with  $\Pi_{B_{1g}}(\Omega) = \sum_j \Pi_j(0, \Omega)$ ]. Obviously, for  $\lambda > 0$  the Raman susceptibility  $\chi_R(\Omega)$  contains no poles and thus orbital fluctuations alone cannot explain the observed resonance in  $\chi_R(\Omega)$ .

- 
- <sup>1</sup> G. R. Boyd, T. P. Devereaux, P. J. Hirschfeld, V. Mishra, and D. J. Scalapino Phys. Rev. B **79**, 174521 (2009); G. R. Boyd, P. J. Hirschfeld, and T. P. Devereaux *ibid.*, 134506 (2010).
- <sup>2</sup> B. Valenzuela, M. J. Calderon, G. Leon, and E. Bascones Phys. Rev. B **87**, 075136 (2013).
- <sup>3</sup> A. V. Chubukov, I. Eremin, and M. M. Korshunov Phys. Rev. B **79**, 220501(R) (2009).
- <sup>4</sup> C. Sauer and G. Blumberg, Phys. Rev. B **82**, 014525 (2010).
- <sup>5</sup> D. J. Scalapino and T. P. Devereaux Phys. Rev. B **80**, 140512(R) (2009).
- <sup>6</sup> S. Sugai, Y. Mizuno, K. Kiho, M. Nakajima, C. H. Lee, A. Iyo, H. Eisaki, and S. Uchida Phys. Rev. B **82**, 140504(R) (2010); *ibid.* **83**, 019903 (2011).
- <sup>7</sup> I. I. Mazin, T. P. Devereaux, J. G. Analytis, Jiun-Haw Chu, I. R. Fisher, B. Muschler, and R. Hackl, Phys. Rev. B **82**, 180502(R) (2010).
- <sup>8</sup> F. Kretzschmar, B. Muschler, T. Bhm, A. Baum, R. Hackl, Hai-Hu Wen, V. Tsurkan, J. Deisenhofer, and A. Loidl, Phys. Rev. Lett. **110**, 187002 (2013).
- <sup>9</sup> V. K. Thorsmølle, M. Khodas, Z. P. Yin, Chenglin Zhang, S. V. Carr, Pengcheng Dai, and G. Blumberg, arXiv:1410.6456.
- <sup>10</sup> W.-L. Zhang, P. Richard, H. Ding, Athena S. Sefat, J. Gillett, Suchitra E. Sebastian, M. Khodas, and G. Blumberg, arXiv:1410.6452.
- <sup>11</sup> Y.-X. Yang, Y. Gallais, R.M. Fernandes, I. Paul, L. Chauvière, M.-A. Méasson, M. Cazayous, A. Sacuto, D. Colson, A. Forget, JPS Conf. Proc. **3** 015001 (2014).
- <sup>12</sup> B. Muschler, W. Prestel, R. Hackl, T. P. Devereaux, J. G. Analytis, Jiun-Haw Chu, and I. R. Fisher, Phys. Rev. B **80**, 180510 (2009).
- <sup>13</sup> Y. Gallais, R. M. Fernandes, I. Paul, L. Chauvière, Y.-X. Yang, M.-A. Méasson, M. Cazayous, A. Sacuto, D. Colson, and A. Forget, Phys. Rev. Lett. **111**, 267001 (2013).
- <sup>14</sup> see, e.g., S. Graser, T. A. Maier, P. J. Hirschfeld, and D. J. Scalapino, New J. Phys. **11**, 025016 (2009).
- <sup>15</sup> M. V. Klein and S. B. Dierker, Phys. Rev. B **29**, 4976 (1984).
- <sup>16</sup> T. P. Devereaux and R. Hackl, Rev. Mod. Phys. **79**, 175 (2007) and references therein.
- <sup>17</sup> T. Strohm and M. Cardona, Phys. Rev. B **55**, 12725 (1997).
- <sup>18</sup> S.L. Cooper and M.V.Klein, Comments Condens. Matter Phys. **15**, 99 (1990); R.R.P.Singh, *ibid.* **15**, 241 (1991).
- <sup>19</sup> R.M. Fernandes, A.V. Chubukov, and J. Schmalian, Nature Phys. **10**, 97 (2014).
- <sup>20</sup> Y. Gallais and I. Paul, C.R. Physique **11**, 113-139 (2016).
- <sup>21</sup> H. Yamase and R. Zeyher Phys. Rev. B **88**, 125120 (2013); New J. Phys. **17**, 073030 (2015).
- <sup>22</sup> G. Blumberg, private communication.
- <sup>23</sup> Y. Gallais, I. Paul, L. Chauvière, and J. Schmalian, arXiv:1504.04570.
- <sup>24</sup> M. Eschrig, Adv. Phys. **55**, 47 (2006).
- <sup>25</sup> Y. Yamakawa, S. Onari, and H. Kontani, arXiv:1509.01161.
- <sup>26</sup> A.V. Chubukov, D. Efremov, I. Eremin, Phys. Rev. B **78**, 134512 (2008); A.V. Chubukov, Phys. C **469**, 640 (2009).
- <sup>27</sup> A. F. Kemper, T. A. Maier, S. Graser, H.-P. Cheng, P. J. Hirschfeld, and D. J. Scalapino, New J. Phys. **12**, 073030 (2010).
- <sup>28</sup> R. M. Fernandes, A. V. Chubukov, J. Knolle, I. Eremin, and J. Schmalian, Phys. Rev. B **85**, 024534 (2012).
- <sup>29</sup> S. Caprara, C. Di Castro, M. Grilli, and D. Suppa Phys. Rev. Lett. **95**, 117004 (2005); see also Hiroshi Kontani,

- Tetsuro Saito, and Seiichiro Onari Phys. Rev. B **84**, 024528 (2011).
- <sup>30</sup> U. Karahasanovic, F. Kretzschmar, T. Bhm, R. Hackl, I. Paul, Y. Gallais, and J. Schmalian, Phys. Rev. B **92**, 075134 (2015).
- <sup>31</sup> M. Khodas and A. Levchenko, Phys. Rev. B **91**, 235119 (2015).
- <sup>32</sup> Ar. Abanov and A.V. Chubukov, Phys. Rev. Lett. **83**, 1652 (1999).
- <sup>33</sup> Chenglin Zhang, H.-F. Li, Yu Song, Yixi Su, Guotai Tan, Tucker Netherton, Caleb Redding, Scott V. Carr, Oleg Sobolev, Astrid Schneidewind, Enrico Faulhaber, L. W. Harriger, Shiliang Li, Xingye Lu, Dao-Xin Yao, Tanmoy Das, A. V. Balatsky, Th. Bruckel, J. W. Lynn, and Pengcheng Dai, Phys. Rev. B **88**, 064504 (2013).
- <sup>34</sup> Q. Q. Ge, Z. R. Ye, M. Xu, Y. Zhang, J. Jiang, B. P. Xie, Y. Song, C. L. Zhang, Pengcheng Dai, and D. L. Feng, Phys. Rev. X **3**, 011020 (2013).
- <sup>35</sup> D. S. Inosov, J. T. Park, P. Bourges, D. L. Sun, Y. Sidis, A. Schneidewind, K. Hradil, D. Haug, C. T. Lin, B. Keimer, and V. Hinkov, Nat. Phys. **6**, 178 - 181 (2010).
- <sup>36</sup> K. Terashima, Y. Sekiba, J. H. Bowen, K. Nakayama, T. Kawahara, T. Sato, P. Richard, Y.-M. Xu, L. J. Li, G. H. Cao, Z.-A. Xu, H. Ding, and T. Takahashi, Proc. Natl. Acad. Sci. U.S.A. **106**, 7330 (2009).
- <sup>37</sup> R. M. Fernandes, L. H. VanBebber, S. Bhattacharya, P. Chandra, V. Keppens, D. Mandrus, M. A. McGuire, B. C. Sales, A. S. Sefat, and J. Schmalian, Phys. Rev. Lett. **105**, 157003; R.M. Fernandes and J. Schmalian, Supercond. Sci. Technol. **25**, 084005 (2012).
- <sup>38</sup> Y. Zhang, C. He, Z. R. Ye, J. Jiang, F. Chen, M. Xu, Q. Q. Ge, B. P. Xie, J. Wei, M. Aeschlimann, X. Y. Cui, M. Shi, J. P. Hu, and D. L. Feng, Phys. Rev. B **85**, 085121 (2012).
- <sup>39</sup> C. He, Y. Zhang, B. P. Xie, X. F. Wang, L. X. Yang, B. Zhou, F. Chen, M. Arita, K. Shimada, H. Namatame, M. Taniguchi, X. H. Chen, J. P. Hu, and D. L. Feng, Phys. Rev. Lett. **105**, 117002 (2010).
- <sup>40</sup> S. T. Cui, S. Y. Zhu, A. F. Wang, S. Kong, S. L. Ju, X. G. Luo, X. H. Chen, G. B. Zhang, and Z. Sun, Phys. Rev. B **86**, 155143 (2012).
- <sup>41</sup> Z.-H. Liu, P. Richard, K. Nakayama, G.-F. Chen, S. Dong, J.-B. He, D.-M. Wang, T.-L. Xia, K. Umezawa, T. Kawahara, S. Souma, T. Sato, T. Takahashi, T. Qian, Yaobo Huang, Nan Xu, Yingbo Shi, H. Ding, and S.-C. Wang, Phys. Rev. B **84**, 064519 (2011).
- <sup>42</sup> I. Paul, Phys. Rev. B **90**, 115102 (2014).
- <sup>43</sup> A.V. Chubukov, Ann. Rev. Condens. Matter Phys. **3**, 57 1686 (2012); A.V. Chubukov, *Itinerant electron scenario for Fe-based superconductors*, Springer Series in Materials Science, (Springer, Berlin, 2015), vol. 211, 255-329.
- <sup>44</sup> Ar. Abanov, A.V. Chubukov, and J. Schmalian, Advances in Physics **52**, 119 (2003).
- <sup>45</sup> S. Avci et al., Nature Commun. **5**, 3845 (2014); A. E. Bohmer, F. Hardy, L. Wang, P. Burger, T. Wolf, P. Schweiss, and C. Meingast, Nat. Commun. **6**, 7911 (2015); X.Wang, J. Kang, and R.M. Fernandes, Phys.Rev.B **91**, 024401 (2015); J. Kang, X. Wang, A. V. Chubukov, and R. M. Fernandes, *ibid* **91**, 121104(R) (2015).
- <sup>46</sup> V. Stanev, J. Kang, and Z. Tesanovic, Phys. Rev. B **78**, 184509 (2008).
- <sup>47</sup> E. Demler, H. Kohno, and S.-C. Zhang, Phys. Rev. B **58**, 5719 (1998); Wei-Cheng Lee, Jairo Sinova, A. A. Burkov, Yogesh Joglekar, and A. H. MacDonald, *ibid.* **77**, 214518 (2008); C. Lee and A.H. MacDonald, *ibid.* **78**, 174506 (2008); I. Fomin, P. Schmitteckert, and P. Wölfle, Phys. Rev. Lett. **69**, 214 (1992); Z. Hao and A. Chubukov, Phys. Rev. B **79**, 224513 (2009).
- <sup>48</sup> A. Hinojosa, A. V. Chubukov, P. Wölfle, Phys. Rev. B **90**, 104509 (2014).
- <sup>49</sup> M. M. Korshunov and I. Eremin, Phys. Rev. B **78**, 140509 (2008); T. A. Maier and D. J. Scalapino, *ibid.* **78**, 020514 (2008).
- <sup>50</sup> V. Cvetkovic and O. Vafek, Phys. Rev. B **88**, 134510 (2013).